

# Document made available under the Patent Cooperation Treaty (PCT)

International application number: PCT/US2005/047572

International filing date: 30 December 2005 (30.12.2005)

Document type: Certified copy of priority document

Document details: Country/Office: US  
Number: 60/643,943  
Filing date: 30 December 2004 (30.12.2004)

Date of receipt at the International Bureau: 20 March 2006 (20.03.2006)

Remark: Priority document submitted or transmitted to the International Bureau in compliance with Rule 17.1(a) or (b)



World Intellectual Property Organization (WIPO) - Geneva, Switzerland  
Organisation Mondiale de la Propriété Intellectuelle (OMPI) - Genève, Suisse

1437522

# THE UNITED STATES OF AMERICA

TO ALL TO WHOM THESE PRESENTS SHALL COME:

UNITED STATES DEPARTMENT OF COMMERCE

United States Patent and Trademark Office

*March 07, 2006*

**THIS IS TO CERTIFY THAT ANNEXED HERETO IS A TRUE COPY FROM THE RECORDS OF THE UNITED STATES PATENT AND TRADEMARK OFFICE OF THOSE PAPERS OF THE BELOW IDENTIFIED PATENT APPLICATION THAT MET THE REQUIREMENTS TO BE GRANTED A FILING DATE.**

**APPLICATION NUMBER: 60/643,943**

**FILING DATE: *December 30, 2004***

**RELATED PCT APPLICATION NUMBER: *PCT/US05/47572***

**THE COUNTRY CODE AND NUMBER OF YOUR PRIORITY APPLICATION, TO BE USED FOR FILING ABROAD UNDER THE PARIS CONVENTION, IS *US60/643,943***



Certified by

Under Secretary of Commerce  
for Intellectual Property  
and Director of the United States  
Patent and Trademark Office

**PRINCETON UNIVERSITY  
OFFICE TECHNOLOGY LICENSING AND INTELLECTUAL PROPERTY  
CONFIDENTIAL INVENTION DISCLOSURE FORM**

Instructions: Please fill out this form as completely as possible. If more space is needed, please attach additional sheets.

**1. Title of Invention:**     **Method of screening compounds that antagonize HAUSP**

Inventor(s) Name(s)*	Princeton Univ. Address	Princeton Telephone	Home Address
Yigong Shi	Department of Molecular Biology 416 Schultz	258-6071	10 Barnyard Court Plainsboro, NJ 08536

**\* The Office of Technology Licensing and Intellectual Property will contact first named inventor to discuss patenting and licensing.**

**2. Category and General Description.** Is the invention a new process, composition of matter, a device, or one or more products? A new use for, or an improvement to, an existing product or process?

The invention can be characterized as a new process. We have made a number of important observations that collectively define novel methods in the identification of compounds that antagonize the function of HAUSP. Because HAUSP is an MDM2-specific deubiquitinating enzyme and MDM2 contributes to several forms of cancer, inhibiting HAUSP may lead to the degradation of MDM2 and subsequent restoration of p53 function.

**3. Utility.** What are possible uses for the invention? In addition to immediate applications, are there any other uses that might be realized in the future?

This invention can be directly used for the identification of novel compounds that inhibit the function of HAUSP.

**4. Novelty.** Pick out and expand on novel and unusual features. How does the invention differ from present technologies currently being used in industry? What problems does it solve, or what advantage does it possess?

p53 is one of the most important tumor suppressor proteins and is inactivated in nearly all cases of cancer. In approximately half of all cancer cases, p53 is inactivated by mutation; in the other half of all cancer cases, wild-type p53 is inactivated by accelerated degradation or mis-localization. MDM2 is an oncoprotein that actively ubiquitinates p53, leading to its degradation. MDM2 also ubiquitinates itself, thus resulting in its own degradation. HAUSP is a member of the UBP family of deubiquitinating

enzymes and selectively targets MDM2 and p53 for deubiquitination, hence rescuing both proteins from degradation. The specific deubiquitination relies on the specific binding of MDM2 and p53 by HAUSP.

We discovered that HAUSP exhibits much higher binding affinity for MDM2 than for p53, implying that MDM2 is the bona fide substrate for HAUSP in cells. More importantly, we discovered that both MDM2 and p53 bind to the same surface pocket in the N-terminal TRAF-like domain of HAUSP. The binding element from MDM2 or p53 is a short peptide composed of no more than 12 amino acids. In addition, we determined a number of crystal structures on HAUSP, including the catalytic core domain, the catalytic core domain bound to ubiquitin aldehyde, the N-terminal TRAF-like domain, and the combined TRAF-like and catalytic core domain. These structures reveal important insights on inhibitor screening. In particular, the TRAF-like domain contains a surface groove that is responsible for binding to the MDM2 or p53 peptide. The catalytic core domain contains a papain-like active site. The C-terminus of ubiquitin is specifically coordinated by a surface cleft in the catalytic core domain.

**5. Method of Synthesis, Assembly, or Process.** If the invention is a composition of matter, a device, or a product, how is it made? If the invention is a process, what are the steps involved?

- (1) Screening for peptide-like small molecules that can bind to the TRAF-like domain of HAUSP;
- (2) Screening for inhibitors that bind to the active site of HAUSP or inhibitors that bind to the surface cleft of HAUSP (this surface cleft specifically binds to the C-terminus of ubiquitin);
- (3) Test the ability of these compounds to inhibit the deubiquitinating activity of HAUSP using polyubiquitinated MDM2 as a substrate;
- (4) Validate lead compounds in cancer cell lines and animal models;
- (5) Improve the efficacy of such compounds and launch clinical trials.

**6. Limitations.** Does this invention possess disadvantages or limitations? Can they be overcome? How?

There is no apparent limitations at present time.

**7. Experimental Verification.** Have you tested the invention experimentally? YES\_\_\_ NO\_\_\_ If yes, describe below. Have you constructed a prototype, model, or test samples which are available for examination? YES\_\_\_ NO\_\_\_ If yes, describe below.

**8. Publications.**

A) Have you described the invention in a publication or an oral presentation? What was the date of publication or presentation? Was it described in specific or general fashion? Include abstracts of talks, news stories, etc. Be sure to include copies of any publications, label them collectively as attachment A.

A manuscript describing this invention will be submitted for publication in the near future. I have not made an oral presentation to the public.

B) Any planned future publications or public disclosure? YES\_\_\_ NO\_\_X\_\_\_ If yes please include a date when the publication will go into the U.S. mail for distribution and the date any public disclosure such as a talk will take place. If there are attachments label as attachment B.

**9. References.** Are there inventions or publications by others that are related to this invention? Please list on a separate sheet, and attach copies if available. Label this list and attach collectively as attachment C.

Maybe. But this invention is very unique and should be distinct from all other inventions.

**10. Sponsorship.** Was the work that led to the invention sponsored? List sponsors below and attach copies of contracts or grant agreements if available. Label these collectively as attachment D.

Sponsor	Contract or Grant Number	Check if copy provided
a) <u>NIH</u>	<u>126-6087</u>	<u>          </u>

**11. MTAs.** Was the invention made with any material or biological substance obtained through a Material Transfer Agreement? Please provide the name of the provider, their affiliation and a copy of the Agreement.

No.

**12. Abstract.** Please provide a brief (less than one page) abstract of the invention. NOTE, THIS ABSTRACT IS INTENDED FOR MARKETING PURPOSES UNDER NONCONFIDENTIAL SITUATIONS. While it is important to include the novel nature of the invention, a general description, intended or proposed uses and utility, the abstract must be a NON-ENABLING DISCLOSURE. Please DO NOT include the essential elements which would allow someone to practice the invention or reproduce the material without a license.

Cancer is one of the most debilitating diseases affecting mankind. More than 600 thousand patients die of cancer in the United States. Oncogenesis is inhibited by the tumor suppressor genes and accelerated by oncogenes. p53 is one of the most important tumor suppressor proteins and is inactivated in nearly all cases of cancer. In approximately half of all cancer cases, p53 is inactivated by mutation; in the other half of all cancer cases, wild-type p53 is inactivated by accelerated degradation or mis-localization, thus unable to perform its function. MDM2 is an oncoprotein that actively ubiquitinates p53, leading to its degradation. MDM2 also ubiquitinates itself, thus resulting in its own degradation. Restoration of p53 function in cancer patients constitutes an important and effective therapeutic method.

HAUSP is a member of the UBP family of deubiquitinating enzymes and selectively targets MDM2 and p53 for deubiquitination, hence rescuing both proteins from degradation. The specific deubiquitination relies on the specific binding of MDM2 and p53 by HAUSP.

We discovered that HAUSP exhibits a much higher binding affinity for MDM2 than for p53, implying that MDM2 is a bona fide substrate for HAUSP in cells. More importantly, we discovered that both MDM2 and p53 bind to the same surface pocket in the N-terminal TRAF-like domain of HAUSP. The binding element from MDM2 or p53 is a short peptide composed of no more than 12 amino acids. In addition, we determined a number of crystal structures on HAUSP, including the catalytic core domain, the catalytic core domain bound to ubiquitin aldehyde, the N-terminal TRAF-like domain, and the combined TRAF-like and catalytic core domain. These structures reveal important insights on inhibitor screening. In particular, the TRAF-like domain contains a surface groove that is responsible for binding to the MDM2 or p53 peptide. The catalytic core domain contains a papain-like active site. The C-terminus of ubiquitin is specifically coordinated by a surface cleft in the catalytic core domain.

Our novel findings immediately identify methods for the screening of anti-HAUSP compounds. The first method involves screening for compounds that bind to the surface groove of the TRAF-like domain. These compounds will prevent the specific recognition of MDM2 by HAUSP, hence accelerating the degradation of MDM2. The second method involves screening for compounds that bind to the surface cleft of the catalytic core domain, which normally coordinates the C-terminus of ubiquitin. These compounds will interfere with the deubiquitinating activity of HAUSP. The third method involves screening for compounds that directly bind to the active site of HAUSP, hence suppressing its catalytic activity.

**13. Possible Means of Commercialization.** How do you envision that the invention might be used in a commercial product or a process for producing a product? Feel free to include as many possibilities as you can. What advantages does this invention have over existing technologies?

Our biochemical and structural information will most likely be explored by pharmaceutical companies for the rational design or screening of drugs. These drugs may be used to treat cancer.

**14. Potential Licensees.** Have you described the invention to industry representatives? Did you describe it in a specific or a general fashion? Did they express any interest? Name them and specific individuals and their titles. Do you know of any other firms that might be particularly interested?

Not yet.

Invention Disclosure Form Completed By: Yigong Shi

Yigong Shi \_\_\_\_\_ December 29, 2004  
Signature Date

Other inventors signatures:

\_\_\_\_\_  
Signature Date

\_\_\_\_\_  
Signature Date

## **HAUSP recognizes multiple substrates through its N-terminal domain**

### **Introduction**

With hundreds of members, the UBP family of deubiquitinating enzymes has become one of the largest enzyme families known to date. For this large family of enzymes, substrate specificity has always been an important yet unresolved issue. The fact that most UBPs contain divergent sequences at their N and/or C terminus strongly indicates that UBPs may have their own specific substrates and the divergent N- or C-terminal sequences are likely to be involved in the

recognition of those substrates. Studies have also shown that, in higher eukaryotes, UBPs are involved in the regulation of various biological processes, including eye development, cell growth, oncogenic transformation and cell cycle regulation. Despite strong evidence suggesting that UBPs may have specialized regulatory functions in multicellular organisms, only very few UBPs with specific substrates have been identified, among which HAUSP is the first mammalian example.

In 1997, HAUSP was first identified as a cellular interacting protein with the herpes simplex virus type I immediate early gene product, Vmw110 (also called ICP0), and was then named HAUSP, which stands for herpes simplex virus associated ubiquitin specific protease. Vmw110 specifically binds to HAUSP C-terminal domain. In addition to Vmw110, another viral protein EBNA1 (Epstein-Barr virus nuclear antigen 1) was also identified to interact with HAUSP. Further studies show that the N-terminal domain of HAUSP is responsible for EBNA1 binding and an EBNA1 peptide of residues 395-450 was sufficient for binding to HAUSP. As will be mentioned later, the HAUSP N-terminal domain, which is responsible for EBNA1 binding, is also responsible for p53 recognition. The fact that EBNA1 binds to HAUSP N-terminal domain and efficiently competes off p53 peptide suggests that EBNA1 might also affect p53 function in vivo by competing for HAUSP.

What makes HAUSP an even more important regulatory protein in cells is the finding that HAUSP can deubiquitinate both p53 and Mdm2, and thus may play a very important regulatory role in the p53-Mdm2 pathway.



The p53 tumor suppressor protein is a sequence-specific transcription factor that can respond to a wide variety of cellular stress signals. The proper function of p53 is indispensable to many cellular processes such as DNA repair, cell cycle arrest, and apoptosis. Loss of p53 function has been observed in almost all tumor types and the gene is mutated in more than 50% of tumors, making it the most commonly mutated gene in human cancer. Mdm2 was later identified to be the major mediator for p53 ubiquitination in normal cells. Mdm2 was first identified as the product of a gene amplified in a spontaneously transformed murine cell line, and was found to be able to inhibit the transcriptional activation activity of p53 by binding to its transactivation domain. Later, it was also found that Mdm2 can promote the ubiquitination and degradation of p53 both in vivo and in vitro. Further studies demonstrated that Mdm2 is a RING finger E3 ubiquitin ligase, and its C-terminal RING finger motif can not only promote p53 ubiquitination both in vivo and in vitro, but also target Mdm2 for autoubiquitination and subsequent degradation. One interesting feature of the Mdm2-p53 regulatory pair is the existence of an autoregulatory feedback loop between Mdm2 and p53. In addition to mediating p53 ubiquitination, Mdm2 is also a negative regulator of p53's transcriptional activity, but on the other hand, the *Mdm2* gene itself is one of the transcriptional targets of p53, so the activation of *Mdm2* gene by p53 would automatically lead to the repression of p53 function.

The Mdm2-p53 pathway is one of the most important pathways involved in p53 regulation. Previous studies imply that multiple mechanisms are involved in

the stabilization of p53 through regulation of the Mdm2-p53 pathway. For example, both phosphorylation and acetylation events can promote p53 stabilization. Furthermore, in response to oncogenic activation, the p14<sup>ARF</sup> tumor suppressor can also stabilize p53 through the inhibition of Mdm2 ubiquitin ligase activity. Besides, the profound effect of ubiquitination on the p53-Mdm2 pathway clearly provides an opportunity for deubiquitination to play a role. Indeed, in year 2002, HAUSP was found to be a novel p53-interacting protein. HAUSP can directly deubiquitinate p53 both in vivo and in vitro, and this deubiquitination function strongly stabilizes p53 and promotes p53-dependent apoptosis.

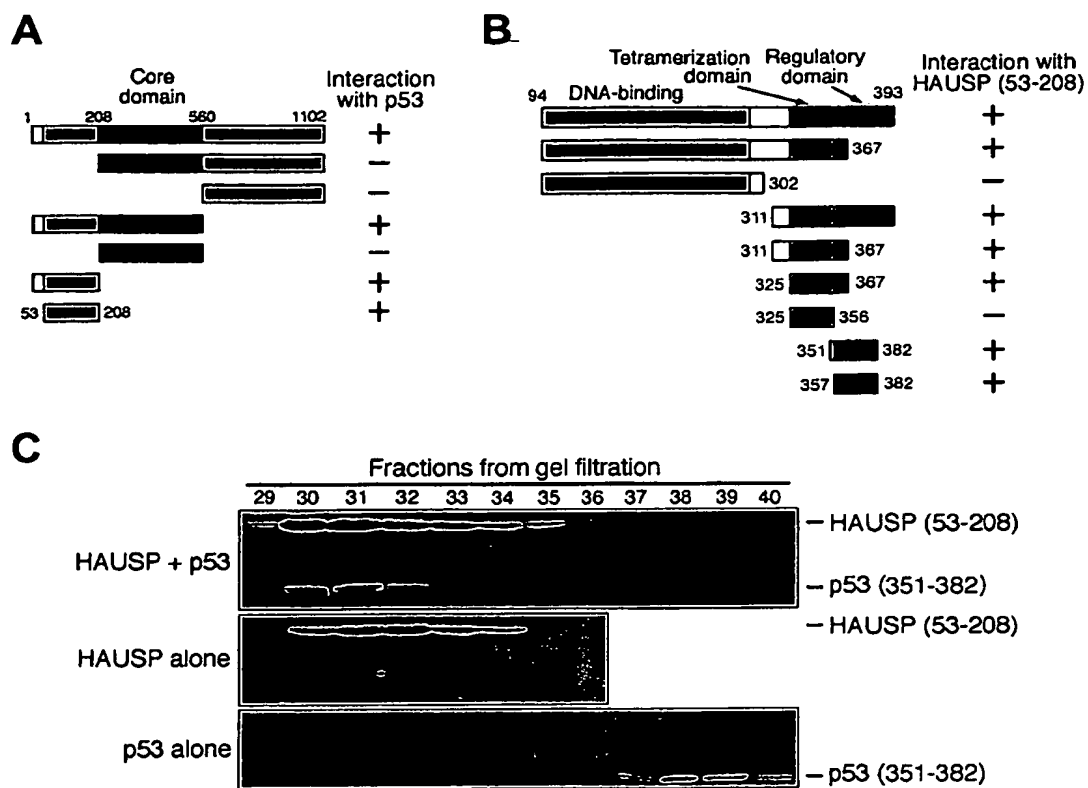
More recently, the picture of HAUSP's function in the p53-MDM2 pathway became more complex with the new discovery that HAUSP exhibits a robust deubiquitinating activity toward Mdm2. Just like p53, MDM2 can also be efficiently deubiquitinated both in vitro and in vivo by HAUSP, and reduction of HAUSP by transient RNAi leads to a decreased half-life in Mdm2. This observation suggests that inactivating HAUSP may have important application in anti-cancer therapies, as it will down-regulate Mdm2. Nonetheless, the exact regulatory mechanisms of HAUSP in the p53-Mdm2 pathway remain incompletely defined, and further structural and/or biochemical characterization of the recognition between HAUSP, p53 and Mdm2 will certainly provide more insights into the real function of HAUSP, and help to elucidate the complex regulatory mechanism in the whole inter-dependent network.

## **Results**

### **HAUSP recognizes the C terminus of p53**

p53 contains an N-terminal transactivation domain (residues 15-30), a central DNA-binding core domain (residues 94-292), a tetramerization domain (residues 326-355), and a C-terminal regulatory domain (residues 356-393). It has been reported that the N-terminal domain (residues 1-248) of HAUSP was sufficient for binding to p53. To further define the domain boundary, we generated a series of deletion variants of HAUSP, expressed and purified these recombinant proteins. We examined their ability to interact with p53 by gel filtration. Based on gel-filtration results, a protease-resistant fragment (residues 53-208) of HAUSP is both necessary and sufficient for stable interactions with p53, while neither the core domain (residues 208-560) nor the C-terminal domain (residues 560-1102) formed a stable complex with p53 (Figure 1).

To identify the minimal sequence requirement in p53, we generated a number of deletion variants and assayed their interaction with HAUSP (residues 53-208) using gel filtration (Figure 1A). Neither the DNA-binding core domain nor the oligomerization domain of p53 is required for the formation of a stable complex with HAUSP. In contrast, a C-terminal 32-residue peptide of p53 (residues 351-382) is sufficient for the interaction with HAUSP (Figures 1B & C). In particular, a short 11-residue peptide stretch (residues 357-367) of p53 plays a critical role in binding to HAUSP as removal of this sequence from p53 (325-356) crippled its interactions with HAUSP (Figure 1B).



**Figure 1** The N-terminal domain of HAUSP recognizes the C-terminal peptide sequences of p53.

**(A)** Identification of the N-terminal domain of HAUSP as the primary p53 binding motif. Various purified recombinant HAUSP fragments were individually incubated with a large p53 fragment (residue 94-393) and then subjected to gel-filtration analysis. The results are summarized.

**(B)** A minimal p53 fragment is both necessary and sufficient for HAUSP binding. p53 peptides with various length and boundaries were purified individually and were tested for binding with the HAUSP N-terminal domain (residues 53-208) by gel-filtration. The results are summarized.

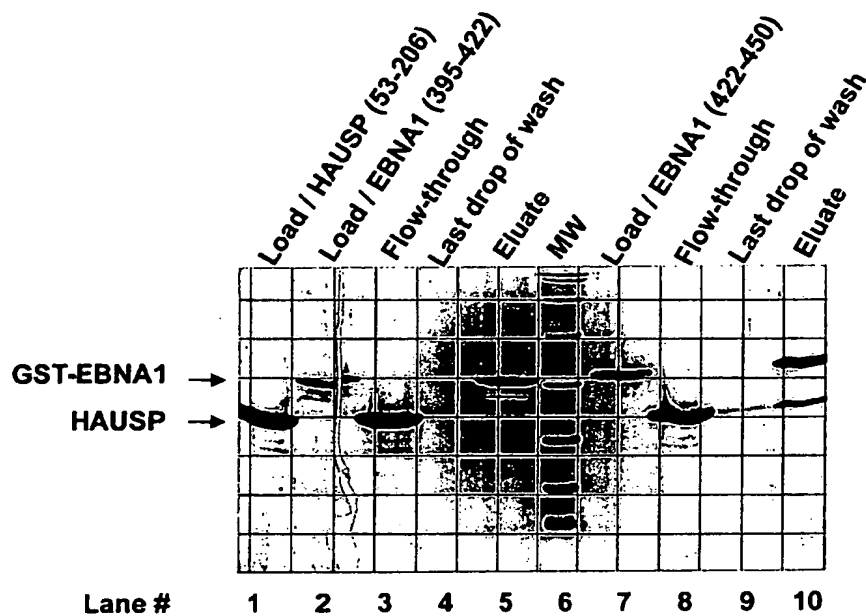
**(C)** A representative gel-filtration run for the complex between HAUSP (residues 53-208) and p53 (residues 351-382). Aliquots of the fractions were visualized by Coomassie staining following SDS-PAGE. The p53 (residues 351-382) peptide, which in isolation was eluted from gel-filtration in fractions 38-39, was eluted in fraction 31-33 when complexed with HAUSP (residues 53-208).

These analyses demonstrate that the N-terminal domain (53-208) of HAUSP stably interacts with a minimal C-terminal peptide (357-382) of p53 (Figure 1B). Supporting this conclusion, the N-terminal domain of HAUSP (residues 58-196) was found to share significant homology (up to 32% sequence identity) to the TRAF (TNF receptor associated factor) domain, a known peptide-binding motif.

### **HAUSP binds to a short peptide motif of EBNA1**

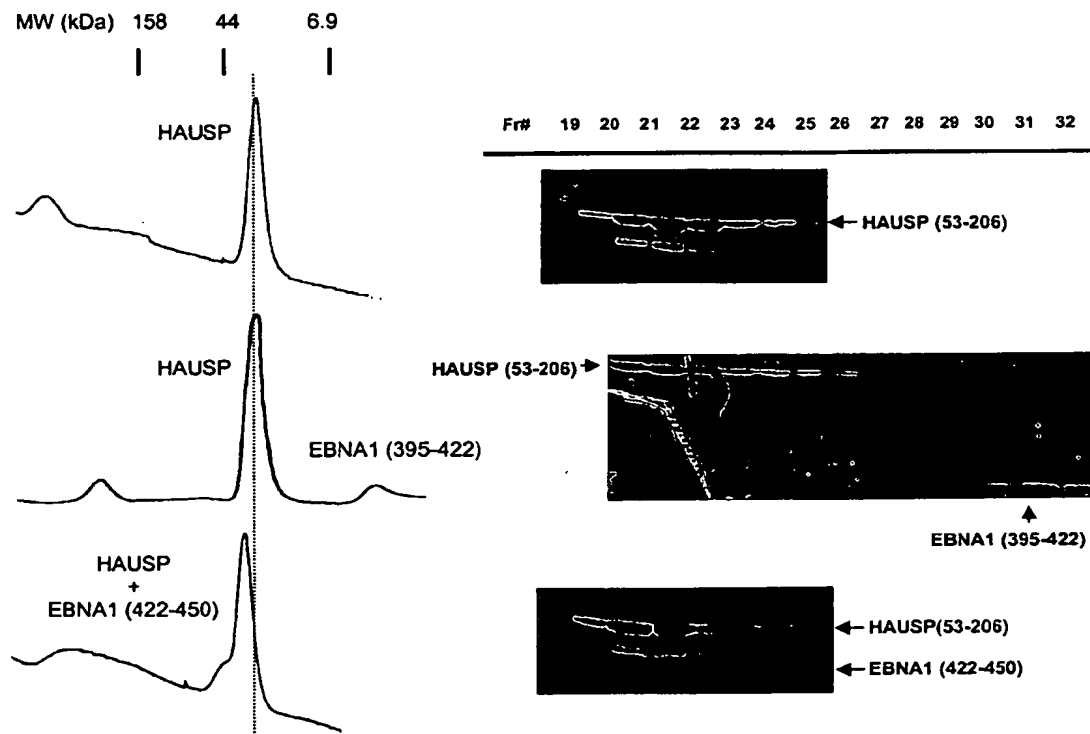
Full-length EBNA1 protein contains 641 amino acids. The identified functional domains of EBNA1 include a long N-terminal Gly-Ala repeat (GAR), a nuclear localization signal, a dimerization domain, and a DNA-binding domain. The GAR in EBNA1 spans a sequence of 238 amino acids (residue 91-328). It helps to block antigen presentation through inhibition of the ubiquitin-dependent proteolysis of EBNA1. A 56 amino-acid region (residues 395-450) in EBNA1 has been shown to be both necessary and sufficient to bind to HAUSP. To further define the HAUSP-interacting sequence in EBNA1, we first generated two deletion variants for the previously identified HAUSP-interacting region (residues 395-450), with one retaining the N-terminal half sequence and the other retaining the C-terminal half sequence. Both deletion variants were purified to homogeneity in their GST-tagged form, and were reloaded to GS4B columns to generate EBNA1 affinity columns. The purified HAUSP N-terminal domain protein (residue 53-206), which was previously shown to interact with EBNA1, was loaded onto the EBNA1 affinity columns and was subjected to extensive

washing afterwards. The binding of EBNA1 deletion variants to HAUSP were assayed through their ability to retain HAUSP N-terminal domain protein on the affinity column. The EBNA1 affinity chromatography results indicate that the N-terminal half of EBNA1 (395-450) doesn't bind to HAUSP N-terminal domain, while the C-terminal half sequence (residues 422-450) is both necessary and sufficient to bind to HAUSP N-terminal domain (Figure 2).



**Figure 2 HAUSP N-terminal domain binds to EBNA1 (422-450) peptide.** EBNA1 affinity columns were constructed by loading purified GST-EBNA1 peptide proteins to GS4B affinity columns (lane 2 and Lane 7), Purified HAUSP N-terminal domain was loaded to both GST-EBNA1 (395-422) and GST-(422-450) columns (Lane 1), and was washed extensively (lane 4 and Lane 9). Protein sample on both columns were eluted off using GSH. The eluate from the GST-EBNA1 (395-422) affinity column contained no HAUSP N-terminal domain protein (Lane 5), while a protein band corresponding to HAUSP N-terminal domain was detected in the eluate from GST-EBNA1 (422-450) column (lane 10).

The results derived from EBNA1 affinity chromatography were further confirmed by gel-filtration chromatography experiment (Figure 3),

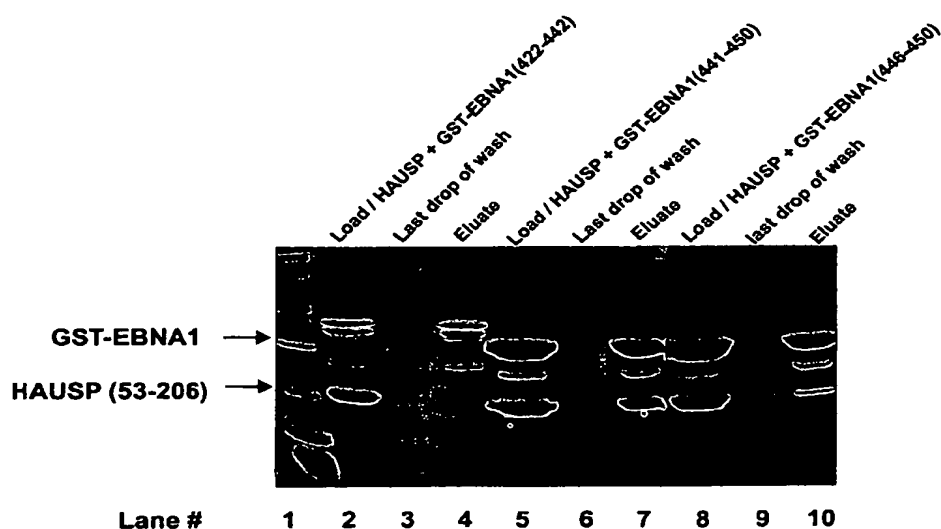


**Figure 3 Gel-filtration runs for the binding test between HAUSP N-terminal domain and EBNA1 peptides.** Left panel shows the gel-filtration chromatography of HAUSP (53-206) by itself (top), HAUSP / EBNA1 (395-422) mixture (middle), and HAUSP / EBNA1 (422-450) mixture. Aliquots of the fractions for each gel-filtration run were visualized by Coomassie staining following SDS-PAGE. Due to the lack of positively-charged residues, EBNA1 (422-450) peptide does not get stained by Coomassie.

According to the gel-filtration results, the EBNA1 (residues 395-422) peptide does not form a stable complex with the HAUSP N-terminal domain in solution, and was eluted separately from HAUSP N-terminal domain on gel-filtration. In contrast, the EBNA1 (residues 422-450) can form a stable complex

with the HAUSP N-terminal domain, and the complex was eluted from gel-filtration one fraction earlier than HAUSP N-terminal domain by itself.

To further delineate the minimal HAUSP-binding sequence in EBNA1, we carried out another round of EBNA1 affinity chromatography to test the HAUSP binding ability of a series of deletion variants within the previously identified HAUSP binding sequence (residues 422-450). In order to avoid potential steric interference from the GST moiety, a flexible linker of 35 amino acids was added in between GST and EBNA1 peptide for each GST-fusion EBNA1 construct. Aliquots of loading samples, wash and eluates were run on SDS-PAGE (Figure 4).



**Figure 4 EBNA1 affinity chromatography.** GST-fusion EBNA1 peptides with different length were loaded to individual GS4B columns. Purified HAUSP N-terminal domain (53-206) was loaded to individual GST-EBNA1 columns (lane 2, lane 5 and lane 8), washed extensively (lane 3, lane 6, and lane 9), and eluted with GSH (lane 4, lane 7, lane 10).



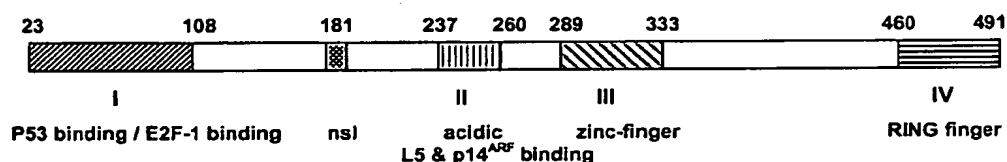
According to the second-round EBNA1 affinity chromatography results, no HAUSP N-terminal domain protein was bound to GST-EBNA1 (422-442), indicating that this part of the sequence may not be involved in HAUSP binding. In contrast, a substantial amount of HAUSP N-terminal domain protein was detected in the eluate from GST-EBNA1 (441-450) column, indicating that this is the most critical sequence for HAUSP binding. A faint band of HAUSP N-terminal domain was also seen in the eluate sample from GST-EBNA1 (446-450), which suggests that there is also binding between HAUSP and EBNA1 (446-450). However, compared to the binding between HAUSP and EBNA1 (441-450), EBNA1 (446-450) binds to HAUSP much less tightly, indicating that further truncation after residue 441 may have removed part of HAUSP binding sequence.

Combining the data from both affinity chromatography and gel-filtration, a minimal sequence of 10 amino acids in EBNA1 (resides 441-450) has been shown to play an essential role in binding to HAUSP N-terminal domain.

### **HAUSP recognizes a central acidic domain of Mdm2**

The full-length human Mdm2 contains 491 amino acids. According to the sequence alignment of Mdm2 from different species, the full-length Mdm2 can be divided into four major conserved regions (Figure 5). Region I is at the N terminus of Mdm2, containing about 90 amino acids without obvious sequence similarity to known domain identities. This region contains the p53-binding domain. Region II is a highly acidic region with 62% sequence homology across

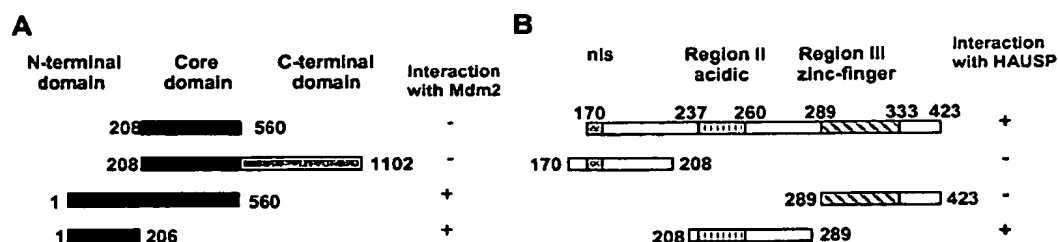
species. This region has been shown to be involved in the binding with multiple proteins, including p14ARF, the inhibitor of Mdm2, and L5, a component of the large ribosomal subunit. Region III is a potential zinc finger domain with 58% sequence homology. The exact function of this domain still remains unclear. Region IV is a highly conserved RING finger domain with 83% sequence homology. This domain is responsible for the E3 ubiquitin ligase activity of Mdm2. A short basic sequence between region I and region II with a potential nuclear localization signal is also highly conserved.



**Figure 5 Schematic representation of the domain compositions of human Mdm2 protein.** Full-length Mdm2 protein contains four major domains (I – IV). The boundary residues for each domain are indicated.

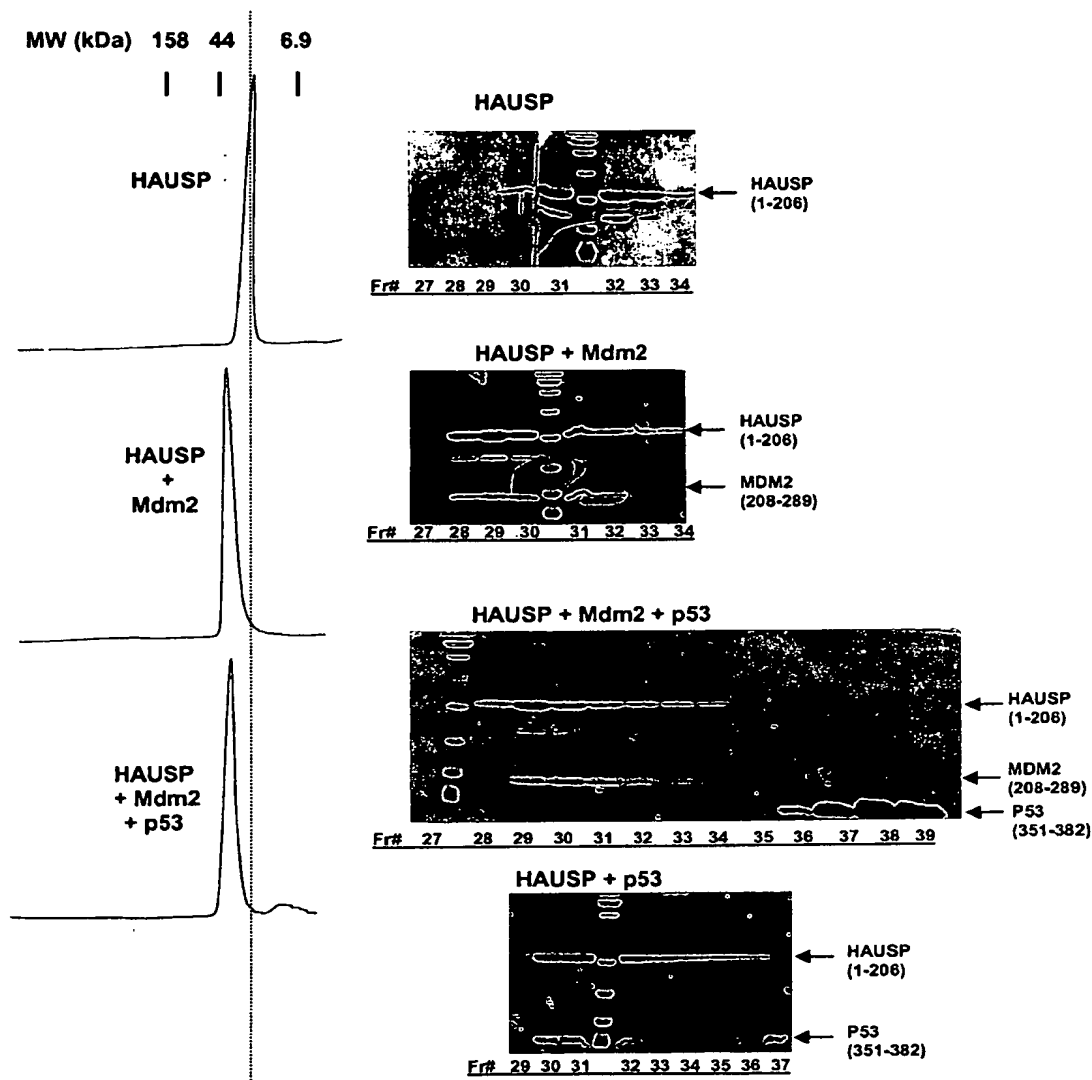
To study the recognition between Mdm2 and HAUSP, we generated a series of deletion constructs for both Mdm2 and HAUSP, expressed and purified these recombinant proteins and tested their ability to interact with each other, using gel filtration. According to gel-filtration results, just as in the case of p53 interaction assay, the N-terminal domain of HAUSP is both necessary and sufficient for stable interactions with Mdm2. Neither the core domain (208-560) nor the C-terminal domain is involved in the binding with Mdm2 (Figure 6A). For

Mdm2, an 82 amino acid sequence (208-289) in the acidic region is sufficient to bind to HAUSP N-terminal domain (Figure 6B).



**Figure 6 The N-terminal domain of HAUSP recognizes the acidic region sequence of Mdm2.** (A) Identification of the N-terminal domain of HAUSP as the primary Mdm2 binding motif. Various purified recombinant HAUSP domain constructs were individually incubated with a large fragment of Mdm2 (residues 170-423) and then subjected to a gel-filtration analysis. The results are summarized. (B) Identification of an Mdm2 fragment including the acidic region (residues 208-289) that is sufficient for HAUSP recognition.

As previously demonstrated, the p53 peptide (residues 351-382) can form a stable complex with the HAUSP N-terminal domain in solution. However, in the presence of Mdm2, the binding between p53 and HAUSP was completely abolished (Figure 7). In the gel-filtration binding assay for HAUSP, Mdm2 and p53, Mdm2 (residues 208-289) efficiently competed off p53 peptide (residues 351-382), and formed a stable complex with HAUSP N-terminal domain (residues 208-289). The currently identified HAUSP recognition motif of Mdm2 includes the central acidic domain of Mdm2. This central acidic domain of Mdm2 contains 37% acidic residues in mouse and human.



**Figure 7 Representative gel-filtration runs for the binding assay between HAUSP and Mdm2.** Left panel shows the gel-filtration chromatography of HAUSP (1-206) by itself (top), HAUSP / Mdm2 (208-289) mixture (middle), and HAUSP / Mdm2 (208-289) / p53 (351-382) mixture (bottom). Aliquots of the fractions for each gel-filtration run were visualized by Coomassie staining following SDS-PAGE. A representative SDS-PAGE gel for the complex between HAUSP and p53 is also shown on bottom right. The HAUSP N-terminal domain (residues 1-206) alone was eluted from gel-filtration with the peak fraction at 32 (1<sup>st</sup> gel, from top), and was eluted with the peak fraction at 29 when complexed with Mdm2 (208-289) (2<sup>nd</sup> gel from top). p53 peptide (residues 351-382) itself can form a complex with HAUSP N-terminal domain (residues 1-206), and the complex was eluted with the peak fraction at 31 (4<sup>th</sup> gel from top). However, in the HAUSP (residues 1-206), Mdm2 (residues 208-289) and p53 (351-382) tertiary mixture, the p53-HAUSP interaction was disrupted and only Mdm2 and HAUSP formed a complex.

It has been reported that this acidic region of Mdm2 also interacts specifically with L5, a component of the large ribosomal subunit, and thus is implicated in ribosome biosynthesis or in translational regulation. In addition to L5 binding, the central acidic region of Mdm2 is essential for ARF binding. ARF (p14ARF and p19 ARF of human and mouse, respectively) is a small basic protein encoded by the INK4a locus, which also encodes the cyclin-dependent kinase inhibitor p16INK4a. Soon after its identification, ARF was found to interact with Mdm2 and to block Mdm2-mediated ubiquitination of p53. A truncated human Mdm2 containing residues 208–491 was able to interact with human ARF, and deletion of residues 222–437 completely abolished its binding to ARF. Further deletion analysis has showed that an acidic domain overlapping region (residues 210-244) of human Mdm2 is most critical for the ARF binding activity. Our finding that the acidic domain is also involved in the recognition with HAUSP adds an additional function to this domain.

#### **HAUSP N-terminal domain adopts a TRAF-like structure**

In addition to the biochemical analyses of the binding between HAUSP and its substrates, a structural approach was used to further study the substrate recognition by HAUSP. Given the fact that the N-terminal domain of HAUSP is involved in the recognition of three out of the four currently identified HAUSP-interacting proteins, the structure of the HAUSP N-terminal domain alone became the first focus. BLAST search results for HAUSP N-terminal domain shows that the HAUSP N-terminal domain shares 34% sequence identity with the

TRAF domain (TD) of human TRAF2. Tumor necrosis factor (TNF) receptor-associated factors (TRAFs) are a family of adapter proteins that were first identified for their ability to bind to TNF family receptors (TNFRs). TRAFs can regulate the functions of the TNFR superfamily by linking the cytosolic domain of those receptors to downstream regulators such as protein kinases or ubiquitin ligases. Up to now, six members of the TRAF family have been identified in human and mice, and two TRAFs have been identified in *Drosophila*. All known TRAFs share a highly conserved region near their C-terminus, which is referred to as the “TRAF domain (TD)”. The TRAF domain represents a novel protein fold of an anti-parallel  $\beta$ -sandwich structure and it usually consists of a sequence of about 180 amino acids. It is well known that TRAFs regulate the function of TNF family receptors mainly through binding to the specific peptidyl motifs in the cytosolic domains of the receptors. The functional specificity of TRAFs arises from the different peptide motifs they recognize.

To elucidate the mechanism of substrate recognition by HAUSP, we crystallized the N-terminal domain of HAUSP (53-206) and determined its structure at 1.6 Å resolution using multi-wavelength anomalous dispersion (Table 1). Consistent with sequence analysis, the HAUSP N-terminal domain adopts a TRAF-like structure (Figure 8). It forms a nine-stranded antiparallel  $\beta$ -sandwich, with strands  $\beta 1$ ,  $\beta 9$ ,  $\beta 5$  and  $\beta 6$  in one sheet and  $\beta 2$ ,  $\beta 3$ ,  $\beta 4$ ,  $\beta 7$  and  $\beta 8$  in the other.

---

Table 1. Summary of Crystallographic Analysis

Data set	Native (HAUSP)	SeMet $\lambda 1$ (peak)	SeMet $\lambda 2$ (inflection)	SeMet $\lambda 3$ (remote)
Beamline	X25	X12C	X12C	X12C
Wavelength (Å)	1.1	0.9794	0.9797	0.95
Resolution (Å)	1.65	2.40	2.40	2.40
Unique reflections	19,080	5801	5768	5667
Data redundancy	6.5	7.1	7.0	7.0
Completeness, %	98.8	90.1	89.5	87.5
(Outer shell)	(92.5)	(48.6)	(47.5)	(40.4)
I/ $\sigma$ (Outer shell)	36.8 (9.72)	15.3 (2.94)	14.4 (2.81)	13.4 (2.27)
R <sub>sym</sub> (Outer shell)	0.047 (0.19)	0.107 (0.31)	0.099 (0.29)	0.115 (0.40)
Anomal. Diff. (%)	n/a	7.0	6.5	7.2
R <sub>Cullis</sub>		0.8	0.71	0.72
Phasing power (centric/acentric)		0.55/0.39	1.16/0.66	0.99/0.49
Overall Figure of Merit(20-2.6 Å):	0.55			

Refinement	HAUSP
Resolution (Å)	500-1.65
Reflections ( F >0)	17,311
All atoms (solvent)	1,335 (174)
R <sub>cryst</sub> /R <sub>free</sub> (%)	19.4/21.2
Rmsd bond length (Å)	0.009 Å
Rmsd bond angle (deg)	1.56

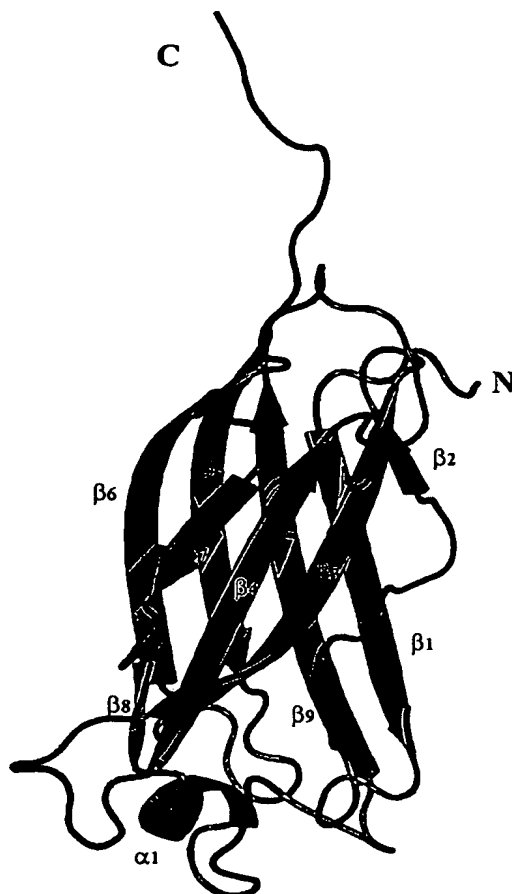
**Ramachandran Plot**

Most favored (%)	86.9
Additionally allowed (%)	13.1
Generously allowed (%)	0
Disallowed (%)	0

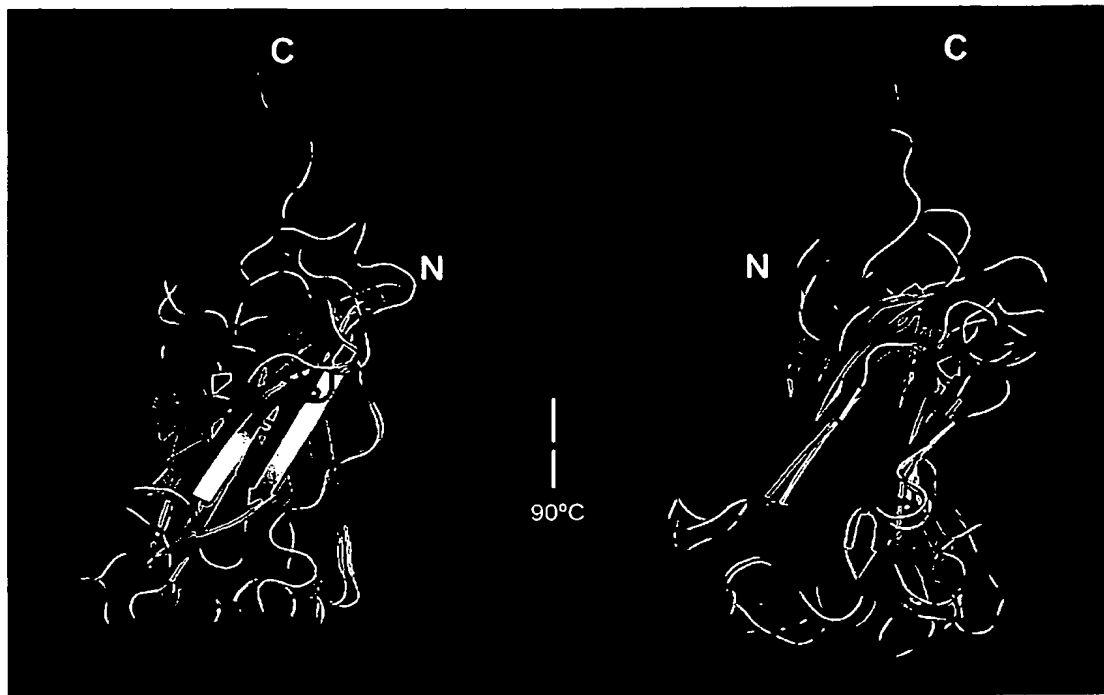
$R_{\text{sym}} = \sum_h \sum_i |I_{h,i} - I_h| / \sum_h \sum_i I_{h,i}$ , where  $I_h$  is the mean intensity of the  $i$  observations of symmetry related reflections of  $h$ .  $R_{\text{cryst}} = \sum |F_{\text{obs}} - F_{\text{calc}}| / \sum F_{\text{obs}}$ , where  $F_{\text{obs}} = F_P$ , and  $F_{\text{calc}}$  is the calculated protein structure factor from the atomic model ( $R_{\text{free}}$  was calculated with 10% of the reflections). Phasing power =  $[(F_{H(\text{calc})}^2 / (F_{PH(\text{obs})} - F_{PH(\text{calc})})^2)]^{1/2}$ , where  $F_{PH(\text{obs})}$  and  $F_{PH(\text{calc})}$  are the observed and calculated derivative structure factors, respectively.  $R_{\text{Cullis}} = \sum ||F_{PH} \pm F_P| - F_{H(\text{calc})}| / \sum |F_{PH} - F_P|$ , where  $F_{H(\text{calc})}$  is the calculated heavy atom structure factor. Figure of Merit =  $\langle \sum P(\alpha) \exp(i\alpha) / \sum P(\alpha) \rangle$ , Where  $P(\alpha)$  is the probability distribution for the phase  $\alpha$ . Rmsd (root-mean-square deviation) in bond lengths and angles are the deviations from ideal values. The different numbers of unique reflections at three different wavelengths were due to slightly different unit cell dimensions.

The topology and overall structure of HAUSP N-terminal domain closely resembles those of the TRAF domain structure. The HAUSP N-terminal domain can be superimposed on the TRAF domain of human TRAF2 structure with an rmsd of 1.48 Å for 101 aligned backbone C $\alpha$  atoms. There is a shallow surface depression in the middle portion of one side of the  $\beta$ -sandwich structure (Figure 9). In TRAF domains, this surface crevice is the region responsible for peptide recognition. In agreement with this feature, our binding assay results have shown that all three substrates known to interact with the HAUSP N-terminal domain use a peptidyl motif for the interaction.





**Figure 8 The overall structure of HAUSP N-terminal domain.** The HAUSP N-terminal domain adopts a TRAF-like structure. It forms a nine-stranded antiparallel sandwich, with strands  $\beta 1$ ,  $\beta 9$ ,  $\beta 5$  and  $\beta 6$  in one sheet and  $\beta 2$ ,  $\beta 3$ ,  $\beta 4$ ,  $\beta 7$  and  $\beta 8$  in the other. A two-turn helix is present in the crossover connection between strands  $\beta 8$  and  $\beta 9$ .



**Figure 9 Structural comparison of HAUSP N-terminal domain with the TRAF domain of human TRAF2.** The HAUSP N-terminal domain (in pink) can be super-imposed to the TRAF domain of human TRAF2 structure (in white) with an rmsd of 1.48 Å for 101 aligned backbone C $\alpha$  atoms. Both proteins adopt an antiparallel  $\beta$  sandwich structures.

## Discussion

Since its identification in 1997, HAUSP has been shown to interact with up to four proteins, each of which has very important regulatory functions in the cell. There are multiple reasons to study the recognition mechanism between HAUSP and its substrates. One arises from the fact that, to date, the issue of substrate specificity of the whole UBP family deubiquitinating enzymes has remained elusive and not well studied. Even though there have been hypotheses claiming that the divergent sequences at the N or C terminus of UBPs endow high substrate specificity, no well-studied example has yet been presented. Being one

of the few UBPs with substrates identified, HAUSP makes a very good target for further studies of the substrate recognition mechanism of the UBP family enzymes. In addition, the fact that all the identified HAUSP-interacting proteins are important regulators in various biological processes makes the study of substrate recognition by HAUSP very important.

Using gel-filtration, we first identified a protease-resistant fragment of HAUSP N-terminal domain (residues 53-208) to be both necessary and sufficient for p53 binding. In a similar approach, we also identified a C-terminal peptide sequence of p53 as necessary and sufficient for binding to the N-terminal p53-recognition motif of HAUSP. The minimal peptide (residues 357-382) comprised only 26 amino acids (Figure 1), but it contains five of the six putative ubiquitination sites. The extensive overlap of these multiple ubiquitination sites with the HAUSP binding site imposes important structural constraints on HAUSP. First, the HAUSP p53-recognition domain must be in close proximity to the active site cleft of the catalytic domain. Second, to accommodate these multiple ubiquitin attachment sites, the connection between the HAUSP catalytic core domain and the p53 binding domain must be flexible in order to position each ubiquitin-p53 linkage within the active site for nucleophilic attack by the catalytic Cys. This organization is likely to be critical for the efficient deubiquitination of p53 in vivo.

The recognition between HAUSP and EBNA1 has been previously characterized. According to literature, EBNA1 binds to the N-terminal domain of HAUSP, the same domain where p53 binds. A 46 amino acid sequence

(residues 395-450) of EBNA1 has been shown to be both necessary and sufficient for HAUSP recognition. In our study, we used a combined approach of both EBNA1 affinity chromatography and gel-filtration to further define the minimal sequence requirement of EBNA1 for HAUSP binding. Based on our results, a short peptide motif of 10 amino acids (residues 441-450) in EBNA1 plays the most critical role for HAUSP binding, and since further deletion of the peptide severely reduced the binding affinity between EBNA1 and HAUSP, it could be the minimal sequence required for HAUSP recognition. The fact that EBNA1 and p53 interact with the same domain of HAUSP raised the possibility that EBNA1 may interfere with the binding between p53 and HAUSP, and indeed this is the case. EBNA1 binds to the N-terminal domain of HAUSP almost 10 times stronger than p53, and the HAUSP-EBNA1 complex formation strongly inhibited the ability of p53 to bind HAUSP. The recognition between p53 and HAUSP has been shown to play an important role in stabilizing p53, and hence inhibiting cell cycle progression and inducing apoptosis. The fact that EBNA1 and p53 interact with the same domain of HAUSP and EBNA1 recognition strongly inhibits p53 binding raised the possibility that EBNA1 may interfere with cellular processes in host cells by disrupting the HAUSP-dependent deubiquitination of p53. By destabilizing p53, EBNA1 would be expected to promote cell cycle progression and prevent apoptosis, which could be important for the host cell immortalization. Epstein-Barr virus efficiently immortalizes cells during its latent infectious cycle, and this process involves a few different Epstein-Barr virus latency proteins. It still remains unclear whether EBNA1 plays a direct role in host

cell immortalization, but the findings that transgenic mice expressing EBNA1 have a tendency to develop B-cell lymphomas suggest that EBNA1 may play a role in those processes. It still remains as a hypothesis that EBNA1 can contribute to host cell immortalization by Epstein-Barr virus through sequestering HAUSP, thereby destabilizing p53. Further studies will be needed to give a clearer answer.

Beside p53 and EBNA1, we also identified that the N-terminal domain of HAUSP is responsible for Mdm2 binding. Due to the difficulty of expressing full-length Mdm2, a long fragment of Mdm2 (residues 170-432) was used to identify the Mdm2-binding region in HAUSP. By gel-filtration assays, we showed that neither the catalytic core domain nor the C-terminal is able to bind to the long fragment of Mdm2. The HAUSP-binding region of Mdm2 was localized to the central acidic region (residues 208-289). Our binding assay results also showed that, like EBNA1, Mdm2 can efficiently compete off p53 peptide from binding to HAUSP. It has been previously demonstrated that HAUSP can efficiently deubiquitinate p53 and Mdm2, both in vivo and in vitro. Based on these findings, a dynamic role of HAUSP in the p53-Mdm2 pathway was proposed, which claims that a kind of substrate balance for HAUSP is achieved under physiological conditions: under certain conditions, HAUSP may exclusively deubiquitinate p53, while under other conditions, Mdm2 may be the major substrate for HAUSP. Our finding that Mdm2 can form a stable complex with HAUSP, and that this complex formation efficiently abolishes p53 binding to HAUSP, strongly argues that in vitro Mdm2 is a preferred substrate for HAUSP. It has been shown that Mdm2 uses

multiple mechanisms to regulate p53 functions, including targeting p53 for ubiquitination and sequestering p53 transactivation activity. Our findings also raised a possibility that in addition to the previously identified regulatory mechanisms, Mdm2 may also be able to destabilize p53 through sequestering HAUSP's deubiquitinating activity on p53.

The HAUSP-binding motif that we identified in Mdm2 includes a sequence motif (residues 210-244) that has been shown to be essential for binding to ARF. ARF protein is previously shown to be a Mdm2 negative regulator. ARF can function to stabilize p53 by blocking the E3 ubiquitin ligase activity of Mdm2. Our finding that the ARF binding motif in Mdm2 overlaps the HAUSP-binding region raised the interesting possibility of an interplay network between Mdm2, HAUSP, and ARF. Further characterization of the interactions between Mdm2 and HAUSP and between Mdm2 and ARF will help to clarify the real function of HAUSP in the ARF-Mdm2 pathway.

As a further effort to study the substrate recognition mechanism of the N-terminal domain of HAUSP, the crystal structure of the HAUSP N-terminal domain was solved. The HAUSP N-terminal domain adopts a nine-stranded antiparallel  $\beta$ -sandwich structure, which closely resembles the TRAF domain (TD) architecture. In the HAUSP N-terminal structure, there is a shallow surface concave at one side of the  $\beta$ -sandwich. In TRAF structures, a similar surface depression is involved in the recognition of specific peptides. This feature strongly supports our conclusion from biochemical binding assays that the nature

of the recognition between HAUSP and its substrates is likely to be peptide-protein interaction.

## Crystal Structure of a UBP-Family Deubiquitinating Enzyme in Isolation and in Complex with Ubiquitin Aldehyde

Min Hu,<sup>1</sup> Pingwei Li,<sup>1</sup> Muyang Li,<sup>2</sup> Wenyu Li,<sup>1</sup>  
Tingting Yao,<sup>3</sup> Jia-Wei Wu,<sup>1</sup> Wei Gu,<sup>2</sup>  
Robert E. Cohen,<sup>3</sup> and Yigong Shi<sup>1,4</sup>

<sup>1</sup>Department of Molecular Biology

Lewis Thomas Laboratory

Princeton University

Princeton, New Jersey 08544

<sup>2</sup>Institute for Cancer Genetics Department  
of Pathology

Columbia University College of Physicians  
and Surgeons

New York, New York 10032

<sup>3</sup>Department of Biochemistry

University of Iowa, Bowen Science Building

51 Newton Road

Iowa City, Iowa 52242

### Summary

The ubiquitin-specific processing protease (UBP) family of deubiquitinating enzymes plays an essential role in numerous cellular processes. HAUSP, a representative UBP, specifically deubiquitinates and hence stabilizes the tumor suppressor protein p53. Here, we report the crystal structures of the 40 kDa catalytic core domain of HAUSP in isolation and in complex with ubiquitin aldehyde. These studies reveal that the UBP deubiquitinating enzymes exhibit a conserved three-domain architecture, comprising Fingers, Palm, and Thumb. The leaving ubiquitin moiety is specifically coordinated by the Fingers, with its C terminus placed in the active site between the Palm and the Thumb. Binding by ubiquitin aldehyde induces a drastic conformational change in the active site that realigns the catalytic triad residues for catalysis.

### Introduction

In eukaryotes, protein ubiquitination is an essential step in the physiological regulation of many cellular processes, such as elimination of damaged or misfolded proteins, cell cycle progression, and signal transduction (Conaway et al., 2002; Glickman and Ciechanover, 2002; Hershko et al., 2000; Hochstrasser, 1996; Laney and Hochstrasser, 1999; Pickart, 2001). Ubiquitin is a highly conserved 76 amino acid polypeptide. Ubiquitin is joined to proteins by an isopeptide bond between the C-terminal carboxylate group of ubiquitin and the lysine  $\epsilon$ -amino group of the acceptor protein. This process depends on the action of three classes of enzymes known as ubiquitin-activating enzyme (E1), ubiquitin-conjugating enzymes (E2), and ubiquitin ligase (E3). Conjugation by ubiquitin is critical for the control of many key regulatory proteins. Accordingly, ubiquitination is itself tightly regulated, and aberrations in this pathway are known to lead to a variety of clinical disorders

(Chung et al., 2001; Schwartz and Ciechanover, 1999).

A major function of ubiquitination is to target proteins for degradation by the 26S proteasome. However, ubiquitinated proteins are not always destroyed by the proteasome. In recent years, protein deubiquitination has emerged as an important regulatory step in the ubiquitin-dependent pathways (Chung and Baek, 1999; D'Andrea and Pellman, 1998; Hochstrasser, 1996; Wilkinson, 1997). Deubiquitination is carried out by the deubiquitinating enzymes (DUBs), which comprise two major groups: the ubiquitin C-terminal hydrolase (UCH) family and the ubiquitin-specific processing protease (UBP) family. Members of the UCH family are usually of small size (~20–30 kDa) and most preferentially cleave ubiquitin from small peptides and amino acids. Unlike the UCHs, the much larger UBPs (60–300 kDa) generally appear to be specific for the proteins targeted for ubiquitination.

The human genome sequencing project led to the identification of more than 90 potential deubiquitinating enzymes, making them one of the largest classes of enzymes in the ubiquitin system. Greater than 80% of these DUBs belong to the UBP class, which appears to regulate a diverse set of biological processes (Chung and Baek, 1999; Hochstrasser, 1996; Wilkinson, 1997). Unlike the highly conserved UCHs, the UBPs contain highly divergent sequences and exhibit strong homology mainly in two regions that surround the catalytic Cys and His residues; these are the so-called Cys Box (~19 amino acids) and the His Box (60–90 amino acids). Although representative UCHs have been structurally characterized (Johnston et al., 1997; Johnston et al., 1999), there is a lack of structural information on the UBPs. Consequently, both the catalytic mechanism and the regulation of the UBP family of DUBs remain unclear.

The p53 tumor suppressor protein is a sequence-specific transcription factor that can respond to a wide variety of cellular stress signals (Levine, 1997; Vogelstein et al., 2000). In an affinity-based approach using GST-p53 as a bait, the cellular protein HAUSP (herpesvirus-associated ubiquitin-specific protease, also known as USP7) was recently identified as a novel p53-interacting protein (Li et al., 2002). HAUSP specifically deubiquitinates the ubiquitinated p53 protein both in vitro and in vivo, and the expression of HAUSP was found to stabilize p53 in vivo and to promote p53-dependent cell growth arrest and apoptosis (Li et al., 2002). These findings reveal an important and novel mechanism in which p53 degradation is prevented by direct deubiquitination and imply that HAUSP might function as a tumor suppressor in vivo through the stabilization of p53. HAUSP is a mammalian UBP for which a specific substrate has been identified.

Here, we report the crystal structures of the 40 kDa catalytic domain of HAUSP in isolation and in a complex with ubiquitin aldehyde. We discuss the structural insights into the catalytic mechanisms as well as the regulation of the UBP family of DUBs. We also present biochemical evidence on how HAUSP recognizes p53.

<sup>4</sup>Correspondence: yshi@molbio.princeton.edu



Table 1. Summary of Crystallographic Analysis

Data set	Native (HAUSP)	SeMet $\lambda 1$ (peak)	SeMet $\lambda 2$ (Inflection)	SeMet $\lambda 3$ (Remote)	HAUSP-Ubal complex
Beamline	Home	X12C	X12C	X12C	X25
Wavelength (Å)	1.5418	0.9787	0.9789	0.9650	1.0056
Resolution (Å)	2.30	2.60	2.60	2.60	2.30
Unique reflections	32,867	25,721	25,507	25,635	64,563
Data redundancy	3.54	7.2	7.4	7.2	4.2
Completeness, % (Outer shell)	96.4 (93.3)	100.0 (100.0)	100.0 (100.0)	100.0 (100.0)	99.7 (99.8)
I/ $\sigma$ (Outer shell)	19.7 (3.3)	25.9 (4.2)	24.4 (4.0)	27.8 (4.8)	16.5 (3.2)
R <sub>sym</sub> (Outer shell)	0.057 (0.37)	0.080 (0.38)	0.077 (0.39)	0.067 (0.30)	0.057 (0.37)
Anomal. Diff. (%)	n/a	6.8	5.8	5.3	n/a
R <sub>cint</sub>		0.53	0.50	0.63	
Phasing power (centric/acentric)		2.88/2.33	3.10/2.52	2.16/1.78	
Overall Figure of Merit (20.0–2.60 Å):	0.63				
Refinement	HAUSP	HAUSP-Ubal complex			
Resolution (Å)	20.0–2.30	25.0–2.30			
Reflections ( F >0)	32,512	59,279			
All atoms (solvent)	5,732 (311)	9,979 (374)			
R <sub>work</sub> /R <sub>free</sub> (%)	22.2/27.9	21.8/26.2			
Rmsd bond length (Å)	0.008 Å	0.008 Å			
Rmsd Bond angle (deg)	1.51	1.46			
<u>Ramachandran Plot</u>					
Most favored (%)	86.0	84.0			
Additionally allowed (%)	12.2	14.7			
Generously allowed (%)	1.3	1.0			
Disallowed (%)	0.5	0.3			

$R_{\text{sym}} = \sum_i \sum_j |I_{ij} - I_i| / \sum_i \sum_j I_{ij}$ , where  $I_i$  is the mean intensity of the  $i$  observations of symmetry related reflections of  $h$ .  $R_{\text{cint}} = \sum |F_{\text{obs}} - F_{\text{calc}}| / \sum F_{\text{obs}}$ , where  $F_{\text{obs}} = F_p$  and  $F_{\text{calc}}$  is the calculated protein structure factor from the atomic model ( $R_{\text{cint}}$  was calculated with 10% of the reflections). Phasing power =  $[(F_{\text{Hickel}}^2 / (F_{\text{Hickel}}^2 - F_{\text{Hickel}}^2))]^{1/2}$ , where  $F_{\text{Hickel}}$  and  $F_{\text{Hickel}}$  are the observed and calculated derivative structure factors, respectively.  $R_{\text{cint}} = \sum |F_{\text{Hickel}} - F_p| / \sum |F_{\text{Hickel}}|$ , where  $F_{\text{Hickel}}$  is the calculated heavy atom structure factor. Figure of Merit =  $\langle \sum P(\alpha) \exp(i\alpha) / \sum P(\alpha) \rangle$ , where  $P(\alpha)$  is the probability distribution for the phase  $\alpha$ . Rmsd (root-mean-square deviation) in bond lengths and angles are the deviations from ideal values. The different numbers of unique reflections at the three different wavelengths were due to slightly different unit cell dimensions.

## Results

### Structure of the Catalytic Core Domain

The full-length HAUSP protein (residues 1–1102) was expressed in baculovirus-infected insect cells and was purified to homogeneity (see Experimental Procedures). HAUSP exists as a monomer in solution as judged by gel filtration and is functional as evidenced by its ability to recognize and deubiquitinate p53 (data not shown). To define the domain boundaries in HAUSP, we employed the approach combining limited proteolysis by subtilisin, sequence alignment with other UBP family members, and binding and deubiquitination assays. This characterization identified a 40 kDa fragment of HAUSP (residues 208–560) as the catalytic core domain, which mediates ubiquitin binding and deubiquitination of substrate.

To gain functional insights into the UBPs, we crystallized the catalytic core domain and determined its structure at 2.3  $\text{\AA}$  resolution using multi-wavelength anomalous dispersion (Table 1 and Figure 1A). There are two molecules of HAUSP in each asymmetric unit; they exhibit an identical set of structural features and can be superimposed with a root-mean-square deviation (rmsd) of 0.6  $\text{\AA}$  for all C $\alpha$  atoms. Thus, we focus our discussion only on one HAUSP molecule (Figures 1B and 2).

The HAUSP catalytic core domain, with a dimension of 75  $\text{\AA} \times 45 \text{\AA} \times 40 \text{\AA}$ , resembles an extended right hand comprised of three domains, Fingers, Palm, and Thumb (Figure 1B). The Thumb consists of eight  $\alpha$  helices ( $\alpha 1$ – $\alpha 6$ ,  $\alpha 9$ , and  $\alpha 10$ ), with the N-terminal Cys Box adopting an extended conformation. The Palm contains eight central  $\beta$  strands ( $\beta 3$ ,  $\beta 8$ – $\beta 14$ ), which are buttressed by two  $\alpha$  helices ( $\alpha 7$  and  $\alpha 8$ ) and several surface loops. An anti-parallel  $\beta$  sheet ( $\beta 8$ ,  $\beta 10$ – $\beta 14$ ), formed by six of the eight  $\beta$  strands from the Palm, intimately packs against the globular Thumb and gives rise to an inter-domain deep cleft. The Cys Box and the His Box are positioned on the opposing sides of this cleft (Figure 1B). The Fingers, comprised of four  $\beta$  strands in the center ( $\beta 1$ ,  $\beta 2$ ,  $\beta 4$ , and  $\beta 7$ ) and two at the tip ( $\beta 5$  and  $\beta 6$ ), reach out more than 30  $\text{\AA}$  from the Palm-Thumb scaffold.

The three-domain structure of HAUSP creates a prominent binding surface between the tip of the Fingers and the Palm-Thumb scaffold. The size and shape of this binding pocket appear to be ideal for the 8 kDa protein ubiquitin (Figure 1B). Moreover, an orientation for the C terminus of a bound ubiquitin is implied by the connection of this binding surface with the deep catalytic cleft between the Palm and the Thumb (Figure 1B, right image). This region of the HAUSP surface is enriched with acidic amino acids.

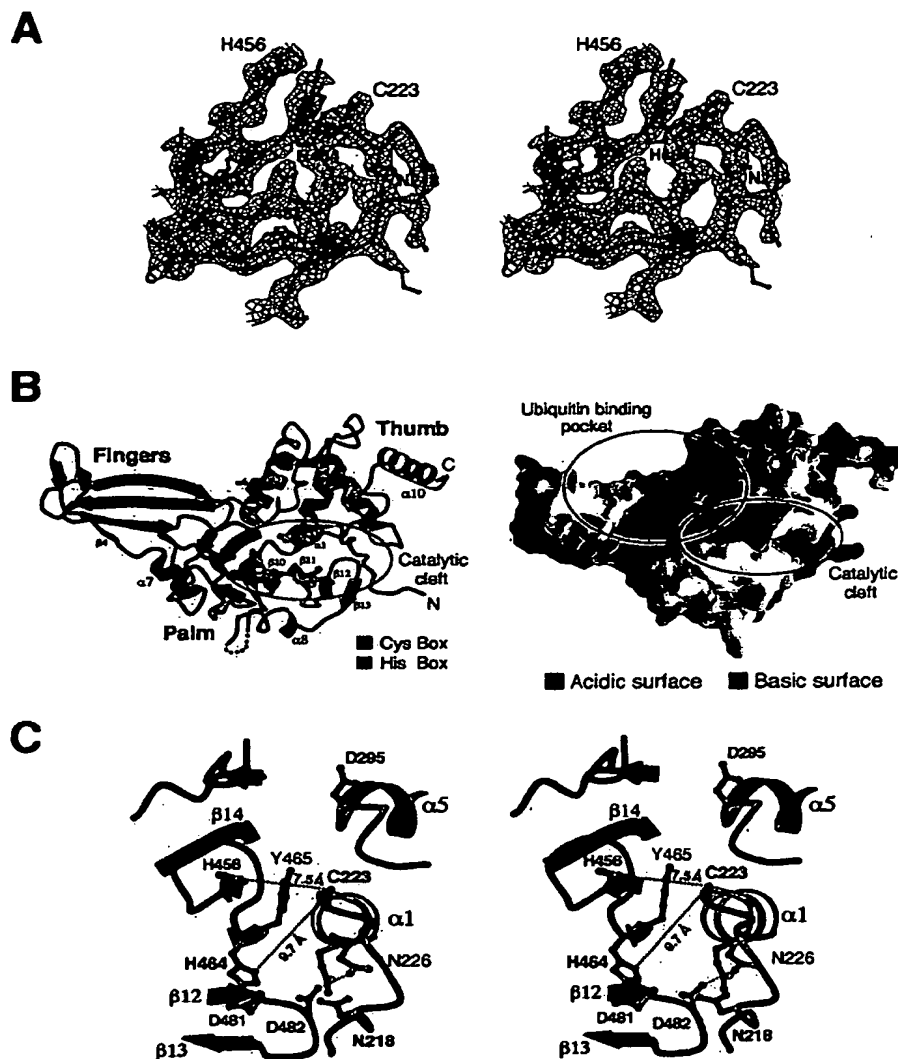


Figure 1. Structure of the Catalytic Core Domain of HAUSP

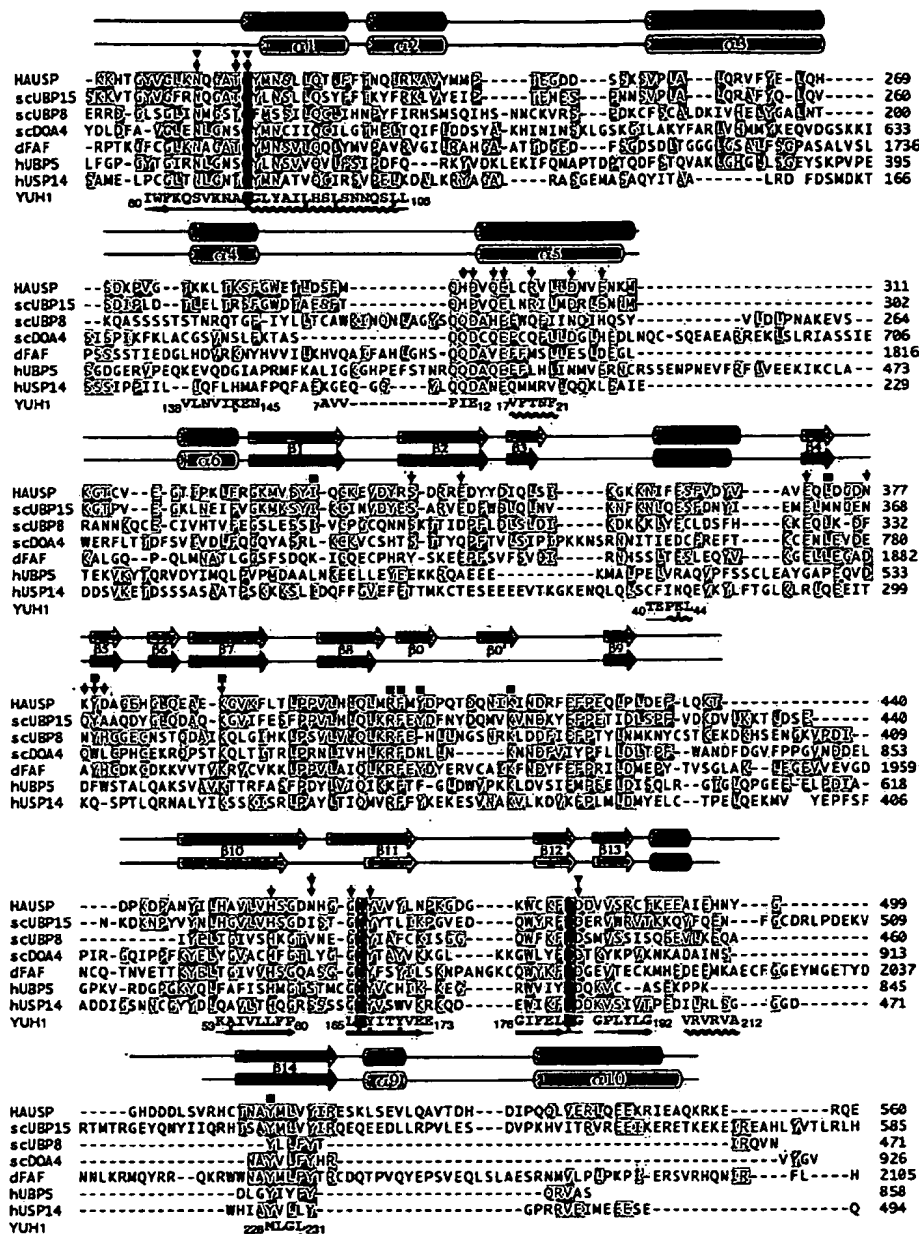
(A) A representative portion of the experimental electron density map in stereo view. The map, contoured at  $1.5\sigma$  at  $2.3\text{ \AA}$  resolution, is shown around the active site region. Some important residues are labeled.

(B) Overall structure of the 40 kDa catalytic core domain of HAUSP. The structure comprises three domains, Fingers (colored green), Palm (blue), and Thumb (gold). The active site, comprising the Cys Box (cyan) and the His Box (purple), is located between the Palm and the Thumb. The surface, represented by electrostatic potential, is shown on the right. A deep cleft runs through the Palm and the Thumb. The predicted ubiquitin binding site is indicated by a black oval circle.

(C) A stereo view of the active site between the Palm and the Thumb. Note that the catalytic residue Cys223 is nearly  $10\text{ \AA}$  away from His464, which was thought to activate Cys223 through deprotonation. All figures were prepared using MOLSCRIPT (Kraulis, 1991) and GRASP (Nicholls et al., 1991).

To investigate whether the three-domain architecture (Fingers-Palm-Thumb) of HAUSP is conserved among other UBP members, we aligned the primary sequences of HAUSP and six representative UBP proteins (Figure 2). The sequence alignment result was crossvalidated with structural information on HAUSP. This analysis revealed that the secondary structural elements in the three domains are generally comprised of conserved

residues (Figure 2). For example, Lys391, on the  $\beta 7$  strand of the Fingers, is invariant among all aligned UBPs. More importantly, residues that contribute to the structural integrity of the Fingers, the Palm, and the Thumb are highly conserved. Notably, the limited hydrophobic core at the tip of the Fingers is formed by the side chains of Tyr379 on strand  $\beta 5$ , Ile332 and Cys334 on strand  $\beta 1$ , and Leu373 on strand  $\beta 4$ . Each of these four



Conserved residues are shaded in yellow whereas the catalytic triad is highlighted in red. Residues that are involved in direct inter-molecular hydrogen bond interactions using their side chains and main chains are marked with purple and green arrows, respectively. Residues that are involved in van der Waals contact with Ubiquitin (Ub) are labeled with blue squares. Residues that coordinate the oxyanion through hydrogen bonds are identified with blue triangles above the alignment. The secondary structural elements above the sequences are indicated for the free HAUSP (lower) and the ubiquitin-bound HAUSP (upper), respectively. The coloring scheme for the secondary structural elements of free HAUSP is the same as in Figure 1. Sequence alignment employed the programs ClustalW (Thompson et al., 1994) and Block Maker (Henikoff et al., 1995). Based on structural similarity, relevant sequences of the UCH protein Yuh1 are also shown below the alignment. Entries shown are from the SwissProt Database: HAUSP (Human; SW:Q93009); UBP15 (*Saccharomyces cerevisiae*; SW:P50101); UBP8 (*S. cerevisiae*; SW:P50102); DOA4 (*S. cerevisiae*; SW:P32571); FAF (*Drosophila melanogaster*; SW:P55824); UBP5 (Human; SW:P45974); USP14 (Human; SW:P54578); and YUH1 (*S. cerevisiae*; SW:P35127).

residues is preserved in at least four other UBP proteins and replaced by a compatible residue in the rest (Figure 2). These observations strongly suggest that the Fingers-Palm-Thumb architecture of HAUSP is conserved among other UBP proteins.

#### Misaligned Active Site of HAUSP

All members of the UBP family of DUBs share strong homology in the Cys and His Boxes. The Cys Box contains the catalytic cysteine residue, which is thought to undergo deprotonation and to unleash a nucleophilic attack on the carbonyl carbon atom of the ubiquitin Gly76 at the scissile peptide bond. In analogy with other cysteine proteases, the deprotonation of this cysteine residue most likely is assisted by an adjacent His residue, which, in turn, is stabilized by a nearby side chain from an Asn or Asp residue. Together, these three residues constitute the so-called catalytic triad. Previous mutagenesis studies on several UBPs have provided evidence that these residues have critical roles in catalysis (e.g., Baek et al., 2001; Huang et al., 1995; Gilchrist and Baker, 2000).

The highly conserved Cys and His Boxes are positioned on opposite sides of the catalytic cleft (Figure 1C). Although Cys223 appears to be the nucleophile in HAUSP, the identities of the other two catalytic residues remain uncertain. His464, a candidate catalytic residue, donates a hydrogen bond to Asp481 (Figure 1C); thus, this His-Asp pair likely constitutes two residues in the catalytic triad. However, the closest distance from the side chain of His464 to that of Cys223 is 9.7 Å, too far away to allow any meaningful interactions. Another invariant histidine residue in the sequence alignment (Figure 2), His456, is separated approximately 7.5 Å away from Cys223 by an intervening residue, Tyr465 (Figure 1C). Moreover, the sulfur atom of Cys223 is not within a reasonable distance of any charged or polar side chain to allow direct interactions. These structural observations indicate that the active site of the free HAUSP core domain exists in an unproductive conformation, and that substrate binding will likely trigger a conformational change that results in catalysis.

#### Overall Structure of a HAUSP-Ubal Complex

Ubiquitin aldehyde (Ubal), a ubiquitin derivative in which the C-terminal carboxylate is replaced by an aldehyde, is a potent inhibitor of most DUBs as it forms a thiohemiacetal with the catalytic Cys, mimicking a reaction intermediate (Hershko and Rose, 1987; Johnston et al., 1999). To elucidate the catalytic mechanisms of the UBPs, we prepared Ubal and reconstituted a covalent complex between HAUSP and the inhibitor. We crystallized this binary complex and solved its structure at 2.3 Å resolution by molecular replacement (Table 1). Each asymmetric unit contains two HAUSP-Ubal complexes; they are nearly identical with a rmsd of 0.53 Å for 419 C $\alpha$  atoms. For simplicity, we focus our discussion on one such complex (Figure 3A).

As anticipated, Ubal binds to the putative substrate binding surface of HAUSP and makes extensive contacts with both the Fingers and the Palm-Thumb scaffold (Figure 3A). The Ubal C terminus is bound in the deep catalytic cleft between the Palm and the Thumb, with a

thiohemiacetal linkage formed between the aldehyde group and the side chain of Cys223. Accompanying the complex formation, two previously flexible surface loops of HAUSP become ordered, with one of these directly contributing to the coordination of Ubal (Figure 3A).

Binding by Ubal induces several prominent conformational changes in the catalytic core domain, resulting in a rmsd of 1.3 Å for 320 aligned C $\alpha$  atoms between the free and Ubal-bound HAUSP structures (Figure 3B). These changes are particularly dramatic in the immediate vicinity of the catalytic cleft, where the main chain loops between  $\alpha$ 4 and  $\alpha$ 5 and between  $\beta$ 10 and  $\beta$ 11 move 5–8 Å (Figure 3B, gold arrows). Interestingly, Ubal binding also draws the tip of the Fingers ( $\beta$ 5) and the edge of the Thumb ( $\alpha$ 5) closer by 2–3 Å (Figure 3B, black arrows); this is likely due to the fact that Ubal makes direct interactions primarily with the tip of the Fingers and the catalytic cleft between the Palm and the Thumb (Figure 3C).

The HAUSP-Ubal interactions result in the burial of 3600 Å<sup>2</sup> exposed surface area (Figure 3C), of which 30% comes from the interface between Ubal and the tip of the Fingers. The rest is derived from the interface between the C-terminal portion of Ubal and its surrounding HAUSP elements, particularly the catalytic cleft. Intriguingly, although some basic residues of Ubal directly interact with the acidic surface of HAUSP, much of the buried surface area on Ubal is uncharged. How can the uncharged Ubal surface interact with the predominantly acidic surface of HAUSP? Examination of the structure reveals that a layer of ordered water molecules are sandwiched between Ubal and the middle portion of the Fingers (Figure 3D). These water molecules form extensive networks of hydrogen bonds among themselves as well as to the side chain and main chain atoms of HAUSP and Ubal. Thus, the binding of Ubal by HAUSP can be visualized as grabbing Ubal with the tip of the Fingers and the catalytic cleft between the Palm and the Thumb, with cushioning waters in between.

These trapped water molecules, absent in the structure of the free HAUSP, exhibit low temperature factors in the complex and might be expected to abate the affinity of HAUSP for ubiquitin and, as a consequence, may enhance the selectivity for ubiquitinated p53 over other ubiquitin conjugates. In support of this analysis, ubiquitin does not form a stable complex with HAUSP as judged by gel filtration. Furthermore, high *K<sub>m</sub>* values were observed in deubiquitination assays with three different substrates (*K<sub>m</sub>*  $\geq$  25  $\mu$ M for ubiquitin-AMC, and *K<sub>m</sub>* > 50  $\mu$ M for ubiquitin-OM or K48-linked diubiquitin; data not shown). These results are in marked contrast to the low micromolar and submicromolar *K<sub>m</sub>* values observed for several UCHs (Larsen et al., 1996; Lam et al., 1997a; Dang et al., 1998; Johnston et al., 1999).

#### Realignment of the Active Site Is Induced by Ubiquitin Binding

The active site of free HAUSP exists in an unproductive conformation (Figure 1C). Upon Ubal binding, structural elements surrounding the catalytic cleft undergo dramatic changes that realign the active site residues for productive catalysis (Figure 4A). Compared to the free HAUSP, Cys223 and His464 shift over a distance of 4.8

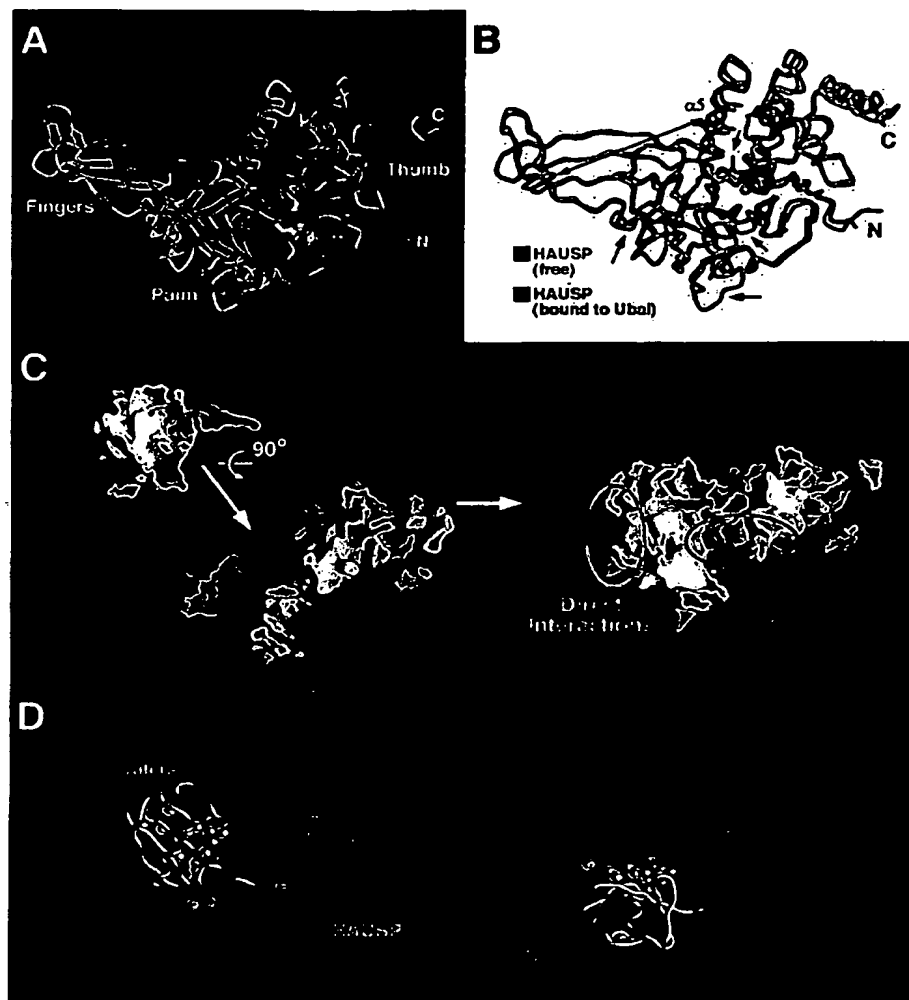


Figure 3. Overall Structure of the HAUSP-Ubal Complex

(A) Overall structure of the catalytic core domain of HAUSP (208–560, blue) covalently bound to Ubal (green). The previously disordered regions in the free HAUSP are highlighted in red. The catalytic triad residues are shown as yellow sticks.

(B) Superposition of structures of HAUSP in isolation (purple) and in complex with Ubal (blue). Red and gold arrows indicate regions that become ordered and that undergo drastic conformational changes upon binding to Ubal, respectively.

(C) Schematic diagram of how ubiquitin is bound by the HAUSP catalytic core domain. After binding, interactions between ubiquitin and HAUSP occur primarily at the opposing ends of the Fingers, as indicated by red circles. Color choices are: blue, the tip of the Fingers; cyan, Cys box; purple, His box; yellow, Ubal C terminus; and green, Ubal surface that interacts with the tip of the Fingers.

(D) Location of buried water molecules. HAUSP is represented as a surface mesh while Ubal is shown as a green coil. A total of 17 buried water molecules are sandwiched between the middle portion of the Fingers and the Ubal. These water molecules mediate numerous hydrogen bonds between HAUSP and Ubal.

and 2.4 Å, respectively, toward the bound C terminus of Ubal (Figure 4A). Consequently, the N<sup>δ1</sup> atom in the imidazole ring of His464 is now 3.6 Å away from the S<sup>γ</sup> atom in the side chain of Cys223, close to a hydrogen bond distance. The stabilizing hydrogen bond between His464 and Asp481 is preserved (Figure 4A). Although the conformational switch involves a distance change of more than 8 Å for regions of the catalytic cleft, there is no disruption of the overall structure and all of the

secondary structural elements are maintained in the Fingers, the Palm, and the Thumb (Figure 2). This remarkable conformational flexibility may be essential to the specific deubiquitination activity of the UBPs.

Another catalytic feature is the formation of the oxyanion hole, which refers to the accommodation of the negative potential formed on the carbonyl oxygen atom at the scissile bond. Typically, the oxyanion is stabilized by hydrogen bonds from the backbone amide group of

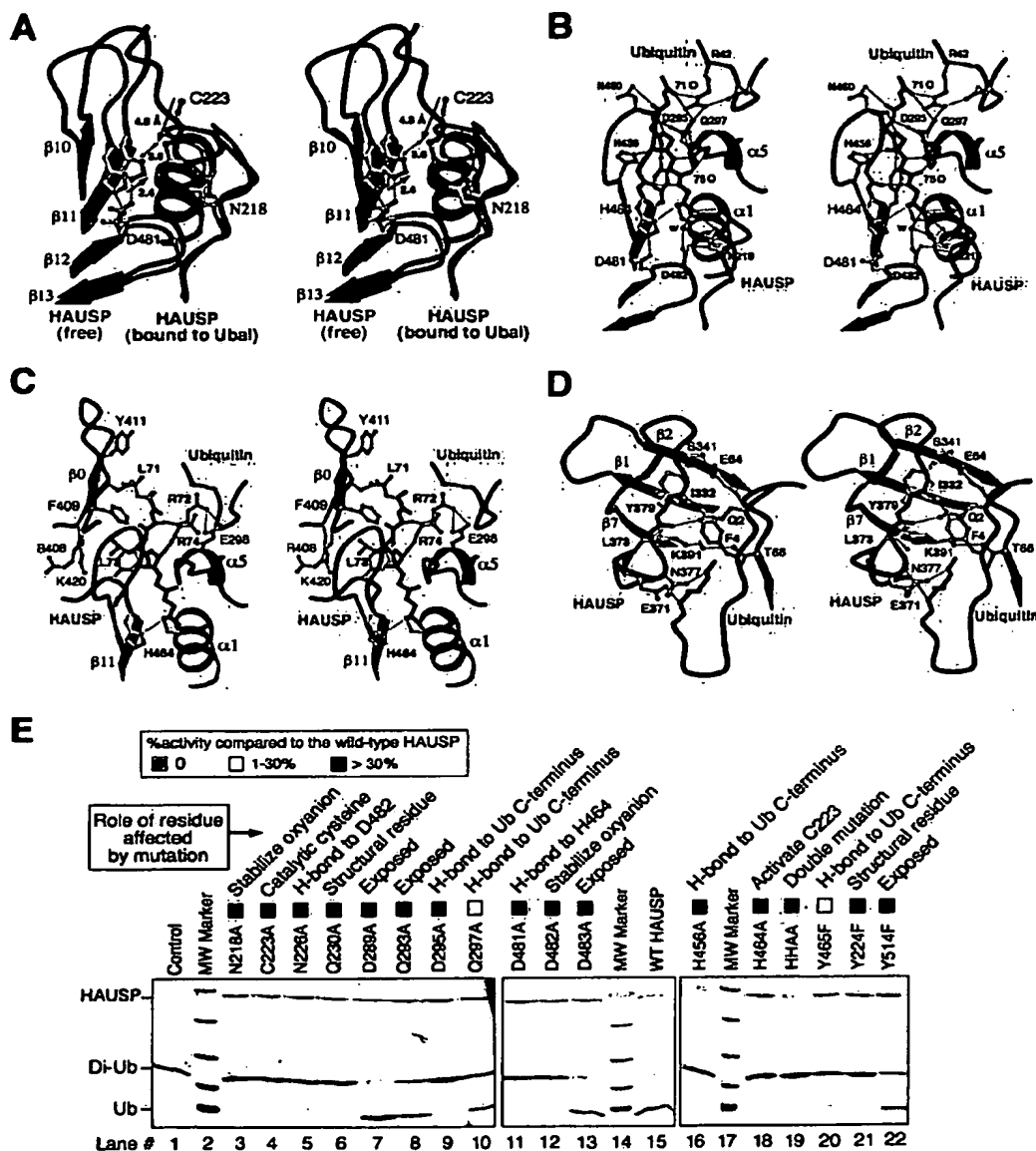


Figure 4. Specific Interactions between HAUSP and Ubal

(A) A large conformational change at the active site induced by Ubal binding. The active sites of HAUSP in isolation (purple) and in complex with Ubal (blue) are superimposed and shown in stereo. The C-terminal tail of Ubal is shown in green. The catalytic triad residues and Asn218 are shown. Note the dramatic conformational changes on all three catalytic residues, which realign these residues for productive catalysis. Hydrogen bonds are represented by red dashed lines.

(B) A stereo view of the hydrogen bonds between HAUSP and the C terminus of Ubal. A water molecule, marked by the letter w, likely plays an important role by hydrogen bonding to the oxyanion and two surrounding residues (Asn218 and Asp482).

(C) A stereo view of the van der Waals interactions between HAUSP and the C terminus of Ubal. The side chains of the Ubal C-terminal six residues as well as several critical HAUSP residues are shown.

(D) A stereo view of the interactions between the tip of the Fingers in HAUSP and Ubal. All HAUSP residues shown in (A–D) are highly conserved among all members of UBPs.

(E) Deubiquitination activity of mutant HAUSP proteins. The substrate used in this assay is Lys48-linked diubiquitin. The role of the affected residue is briefly indicated. HHAA represents the double mutation H456A/H464A. Mutation of any catalytically important residue leads to complete abolishment of the deubiquitination activity.

the catalytic cysteine as well as from a neighboring Gln or Asn. It was previously unclear how the oxyanion is coordinated in UBPs. In the HAUSP-Ubal complex, the oxyanion that normally would form during catalysis is mimicked by the hydroxyl of the thiohemiacetal (i.e., the oxygen from Ubal residue 76). This group can accept two hydrogen bonds and is within hydrogen bond distance of three potential donors, the backbone amide of Cys223 and Thr222 and the side chain of Asn218 (Figure 4B). In addition, a nearby water molecule, supported by two direct hydrogen bonds from Asp482 and Asn218, further neutralizes the negative potential of the oxyanion (Figure 4B). This network of hydrogen bonds is further buttressed by an additional intra-molecular contact between the side chains of Asp482 and Asn226.

Significantly, residues that are involved in the coordination of the oxyanion are highly conserved among all members of the UBPs (Figure 2). For example, Asn218 and Asp482 are invariant among the seven representative UBP proteins whereas Asn226 is present in six members and is replaced by Ser in UBP8 (Figure 2). This analysis suggests a conserved mechanism for all UBPs in the formation of the oxyanion hole and catalysis. The active site geometry of the Ubal-bound HAUSP closely resembles that of the papain family of cysteine proteases (Rawlings and Barrett, 1994).

#### Specific Interaction between HAUSP and Ubal

Ubal primarily contacts two regions of HAUSP, the tip of the Fingers and the catalytic cleft region between the Palm and the Thumb (Figure 3C). The C terminus of Ubal is primarily coordinated by a network of 12 hydrogen bonds between the main chain groups of Ubal and the main chain and side chain atoms of HAUSP (Figure 4B). Notably, all backbone carbonyl carbon and amide nitrogen atoms in the C-terminal five residues of Ubal are involved in direct hydrogen bonds to neighboring residues in HAUSP (Figure 4B).

In addition to hydrogen bonds, inter-molecular van der Waals contacts also contribute to the specific recognition of the C terminus of Ubal. The two hydrophobic residues in the C terminus of Ubal, Leu71 and Leu73, point into the catalytic cleft by making direct van der Waals contacts to the surrounding HAUSP residues (Phe409 and Tyr411 on the newly formed strand  $\beta$ 0, Lys420 on strand  $\beta$ 0', and Tyr514 on strand  $\beta$ 14) (Figures 2 and 4C). In contrast, the two positively charged residues in the C terminus of Ubal, Arg72 and Arg74, point up into the open space away from the cleft. Only Arg72 hydrogen bonds to Glu298 of HAUSP. The last two residues of Ubal, Gly75 and Gly76, fit in the narrowest region of the catalytic cleft between the Palm and the Thumb. The space surrounding the backbone C $\alpha$  atoms of these two glycine residues is insufficient to accommodate any other side chain, consistent with the specific function of the UBPs.

On the other side of the HAUSP-Ubal interface, residues at the tip of the Fingers also make important contributions to the binding of Ubal. These interactions, including 12 direct hydrogen bonds and one patch of van der Waals contacts (Figure 2), primarily involve the N-terminal residues of Ubal as well as a few amino acids in the middle stretch. Gln2, Lys48, Glu64, and Thr66

make two hydrogen bonds each to the HAUSP residues Lys378 and Asp380, Asp305 and Glu308, Ser341 and Tyr379, and Glu345 and Lys391, respectively (Figures 2 and 4D). In this interface, the only van der Waals interactions occur between Phe4 of Ubal and a hydrophobic pocket formed by Ile332, Leu373, Tyr379, and the aliphatic portion of the side chain of Lys391 in HAUSP (Figure 4D).

All of the important HAUSP residues responsible for binding to Ubal are conserved among the seven aligned UBP members (Figure 2). Out of a total of 25 invariant residues among these proteins, 14 contribute directly to the specific recognition of ubiquitin. For example, Asp295 and Glu298 on helix  $\alpha$ 5, His456 on strand  $\beta$ 10, and Tyr465 on strand  $\beta$ 11, all invariant among the seven UBP proteins (Figure 2), make direct hydrogen bonds to coordinate the C terminus of Ubal. Lys391, an invariant residue on strand  $\beta$ 7 of the Fingers, mediates both direct hydrogen bond (to Thr66) as well as van der Waals contacts (to Phe4 of Ubal). The invariant Phe409 and highly conserved Arg408 form part of the hydrophobic pocket to hold Leu71 and Leu73 of Ubal. This analysis solidifies the notion that all UBPs share conserved mechanisms of ubiquitin binding and catalysis. In addition, these observations further support the generality of the observed Fingers-Palm-Thumb architecture of the UBPs.

#### Mutational Analysis

To corroborate our structural studies, we assayed deubiquitination *in vitro* using Lys48-linked diubiquitin as the substrate. Quantitative assays (see Experimental Procedures) with subsaturating substrate (i.e.,  $[S] \ll K_m$ ) established that the HAUSP catalytic domain and full-length HAUSP cleave diubiquitin with similar rates ( $k_{cat}/K_m = 13$  and  $18 \text{ min}^{-1} \text{ mM}^{-1}$ , respectively). To probe the functional significance of the residues in the catalytic cleft, we performed alanine-scanning mutagenesis on 29 amino acids in the Palm and the Thumb domains of the HAUSP catalytic domain. All 29 mutant proteins were purified to homogeneity and were individually examined for their activity in a semiquantitative *in vitro* assay; some relevant results are shown (Figure 4E). As anticipated, whereas the wild-type enzyme was able to reduce the substrate to mono-ubiquitin, mutation of any of the catalytic triad residues (Cys223, His464, and Asp481) or residues that comprise the oxyanion hole (Asn218, Asn226, and Asp482) resulted in an undetectable level of catalytic activity (Figure 4E). In addition, mutation of His456 or Asp295, which hydrogen bonds to the carbonyl of Arg74 or the amide of Leu73 of Ubal, respectively, also led to undetectable activity. In contrast, mutation of any of the four solvent-exposed residues (Asp289, Gln293, Asp483, and Tyr514) had little effect on the deubiquitination activity of HAUSP (Figure 4E). Two residues, Gln230 and Tyr224, contribute to the structural integrity of HAUSP by making contacts to surrounding residues; their mutations also resulted in loss of activity (Figure 4E). These results are in complete agreement with our structural analysis.

#### Structural Comparison with UCH and ULP

To reveal common features of catalysis and substrate binding, we superimposed the structure of the HAUSP-

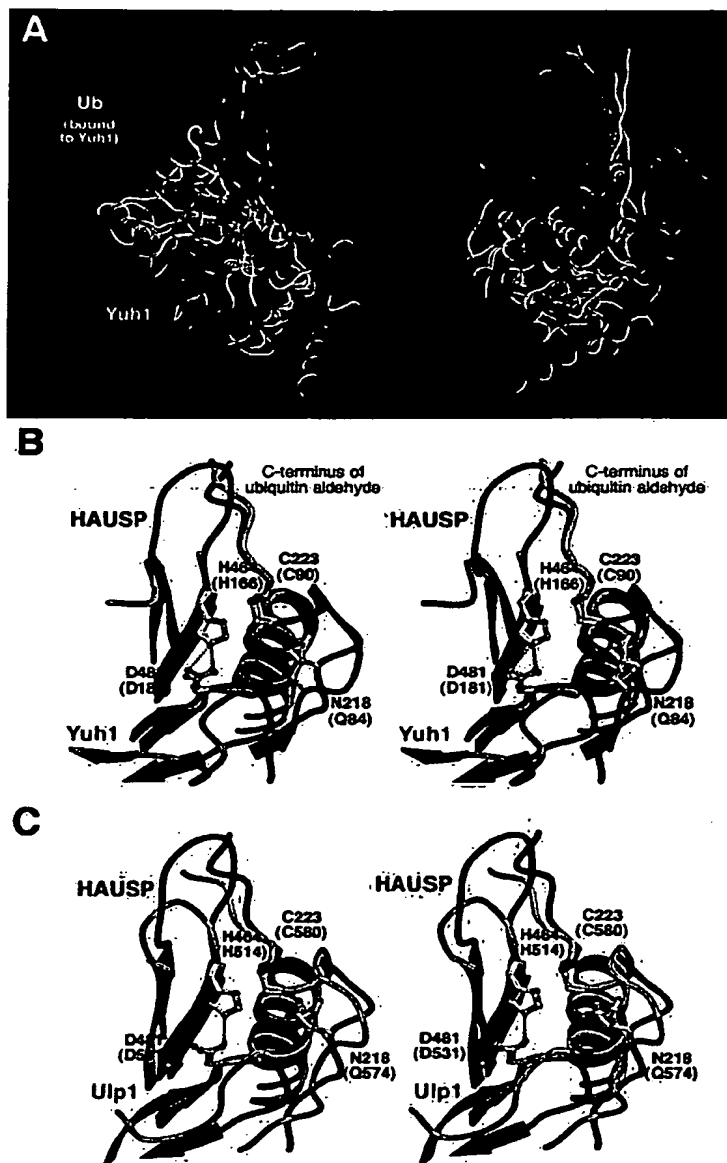


Figure 5. Structural Comparison of HAUSP-Ubal with Yuh1-Ubal and Ulp-Smt3 Complexes

(A) Superposition of the structure of the HAUSP-Ubal complex with that of Yuh1-Ubal. These two structures are aligned on their deubiquitination domains. HAUSP and Yuh1 are shown in blue and purple, respectively. The HAUSP- and Yuh1-associated Ubal moieties are represented as transparent surfaces colored green and yellow, respectively. Two perpendicular views are shown.

(B) Stereo comparison of the active sites of HAUSP (blue) and Yuh1 (purple). The C-terminal tails of Ubal is shown in green and yellow for the HAUSP and Yuh1 complexes, respectively. Catalytic triad residues and the oxyanion-coordinating residue are shown. Hydrogen bonds are represented by red dashed lines. The catalytic triad residues are superimposed with a rmsd of 0.23 Å.

(C) Stereo comparison of the active sites of HAUSP (blue) and Ulp1 (gold). The C-terminal tails of Ubal and Smt3 (SUMO homolog in yeast) are shown in green and yellow, respectively. Catalytic triad residues and the oxyanion-coordinating residue are shown.

Ubal complex onto that of the Yuh1-Ubal complex (Johnston et al., 1999). This alignment resulted in a rmsd of 1.87 Å for 96 similar backbone C $\alpha$  atoms of the deubiquitination domain, with 12% sequence identity (Figures 5A and 2). Both the topology and the overall structure of HAUSP are quite different from that of Yuh1, a member of the UCH family of DUBs. Due to a different topology, sequences of the Yuh1 protein are permuted in this alignment (Figure 2). Yuh1 lacks the Fingers domain. In addition, the 8- $\alpha$  helix Thumb domain of HAUSP is reduced to a tight 5-helical bundle in Yuh1 (Figure 5A). Although the Palm domain of HAUSP maintains the same overall fold as the corresponding regions in Yuh1,

large deviations are apparent throughout the structure (Figure 5A).

Recognition of Ubal is also quite different between HAUSP and Yuh1. Due to the lack of the Fingers, Yuh1 has only one contiguous interface that contacts Ubal. Although the side chain of Arg74 anchors the C terminus of Ubal by making a total of six hydrogen bonds to Yuh1, the same group has no interaction with surrounding residues in HAUSP. In contrast, although Ubal Leu71 contributes little to Ubal-Yuh1 interactions, this residue makes multiple van der Waals contacts to two aromatic residues in HAUSP, Phe409 and Tyr 411.

Despite divergent structures, HAUSP and Yuh1 ex-



hibit a nearly identical geometry at the active site (Figure 5B). The catalytic triad residues of HAUSP can be superimposed with those of Yuh1 with a rmsd of 0.23 Å. The oxyanion is also partially stabilized by similar residues in the same general area, Asn218 in HAUSP and Gln84 in Yuh1. This analysis indicates that hydrolysis of bound substrates by the UCHs and the UBPs is likely to employ the same mechanism. However, important differences in the active site geometry exist prior to substrate binding. For Yuh1, the active site residues are unlikely to undergo any significant conformational change during catalysis as the Ubal-bound conformation is nearly identical to that in the unliganded structure of another UCH member, UCH-L3 (Johnston et al., 1997). In the case of HAUSP, a dramatic structural shift involving the active site residues takes place upon Ubal binding (Figure 4A).

The surprisingly large differences between the structures of the UBPs and the UCHs may be related to their functional differences in biology. For example, the UCHs contain an active site crossover loop that does not allow passage of large substrates such as folded proteins (Johnston et al., 1999). This structural feature indicates that, if unassisted, the UCHs are unlikely to deubiquitinate proteins. On the other hand, the open-cleft structure of the HAUSP catalytic domain and the unique three-domain architecture are ideal for the deubiquitination of large substrates such as a polyubiquitin chain or ubiquitinated protein.

In addition to ubiquitination, conjugation of the small ubiquitin-related modifier (SUMO) to proteins also regulates numerous cellular processes (Hochstrasser, 1998). The crystal structure of an ubiquitin-like protein specific protease (ULP), Ulp1, which catalyzes the deconjugation of SUMO (Smt3 in yeast) from target proteins, was reported in a covalent complex with Smt3 (Mossessova and Lima, 2000). Although the overall structure of the Ulp1-Smt3 complex is quite different from the HAUSP-Ubal complex with 1.7 Å rmsd for 62 aligned C $\alpha$  atoms (data not shown), the active site conformation is highly conserved (Figure 5C). Similar to Yuh1, the catalytic triad of Ulp1 can be superimposed with that of HAUSP with 0.21 Å rmsd. In addition, the stabilization of the oxyanion involves a similar residue (Gln574) in a similar location in Ulp1 (Figure 5C). These analyses indicate that the UBPs, the UCHs, and the ULPs all employ a highly conserved catalytic mechanism for the deconjugation of ubiquitin and SUMO from target proteins.

#### HAUSP Recognizes the C Terminus of p53

p53 contains an N-terminal transactivation domain (15–30), a central DNA binding core domain (94–292), a tetramerization domain (326–355), and a C-terminal regulatory domain (356–393) (Levine, 1997). It has been reported that the N-terminal domain (residues 1–248) of HAUSP was sufficient for binding to p53 (Li et al., 2002). The N-terminal 52 residues of HAUSP are 75% hydrophilic and are likely to be flexible in solution as they are readily removed from the rest of the sequences by limited proteolysis. To further define the domain boundary, we generated a series of deletion variants of HAUSP, expressed and purified these recombinant proteins, and examined their ability to interact with p53. Using gel filtration, we found that a protease-resistant fragment

(53–208) of HAUSP was both necessary and sufficient for stable interactions with p53 whereas neither the core domain (208–560) nor the C-terminal domain (560–1102) formed a stable complex with p53 (Figure 6A).

To identify the minimal sequence requirement in p53, we generated a number of deletion variants and assayed their interaction with HAUSP (53–208) using gel filtration (Figure 6B). Neither the DNA binding core domain nor the oligomerization domain of p53 was required for the formation of a stable complex with HAUSP. Rather, a C-terminal 32-residue peptide of p53 (351–382) contained the determinants for the interaction with HAUSP (Figures 6B and 6C). In particular, a short 11-residue peptide stretch (357–367) of p53 plays a critical role in binding to HAUSP as removal of this sequence to generate p53 (325–356) eliminated the interaction with HAUSP (Figure 6B). These analyses demonstrate that the N-terminal domain (53–208) of HAUSP stably interacts with a minimal C-terminal peptide (357–382) of p53 (Figure 6B). Supporting this conclusion, the N-terminal domain of HAUSP (residues 58–196) was found to share significant homology (up to 32% sequence identity) to the TRAF (TNF receptor associated factor) domain (Zapata et al., 2001), a known peptide binding motif (Chung et al., 2002).

#### Discussion

Although protein ubiquitination and subsequent proteasomal degradation have been studied extensively, protein deubiquitination and hence stabilization remain less well characterized (Chung and Baek, 1999; D'Andrea and Pellman, 1998; Wilkinson, 1997). Many and possibly most of the UBPs function to deubiquitinate conjugates of specific proteins. HAUSP is a UBP originally identified as a cellular protein that interacts with the immediate-early protein Vmw110 of the Herpes simplex virus (Everett et al., 1997). Recently, HAUSP was found to modulate the stability and function of the tumor suppressor protein p53 via specific deubiquitination (Li et al., 2002). In this study, we report the structure of a UBP catalytic core domain, which reveals a novel three-domain architecture comprised of the Fingers, Palm, and Thumb domains. It is of particular note that this nomenclature should not be confused with that of the DNA polymerase, where Fingers, Palm, and Thumb refer to different entities (Steitz, 1999 and references therein). We also report the structure of the catalytic core domain of HAUSP bound to ubiquitin aldehyde, which reveals a dramatic conformational change in the active site upon binding by Ubal. The conformational state of HAUSP bound to Ubal is likely to resemble that of HAUSP bound to substrate.

In a search for structural homologs using the program DALI (Holm and Sander, 1993), the Ca<sup>2+</sup>-bound form of the  $\mu$  calpain protease core (PDB code 1KXR) was found to exhibit the highest degree of similarity to HAUSP, with a rmsd of 2.1 Å for 87 aligned C $\alpha$  atoms (9% sequence identity). Interestingly, Ca<sup>2+</sup> binding to  $\mu$  calpain also induces a large conformational switch in the active site, realigning a separated catalytic triad into a productive conformation (Moldoveanu et al., 2002). The extent of this change upon Ca<sup>2+</sup> binding is similar to that observed

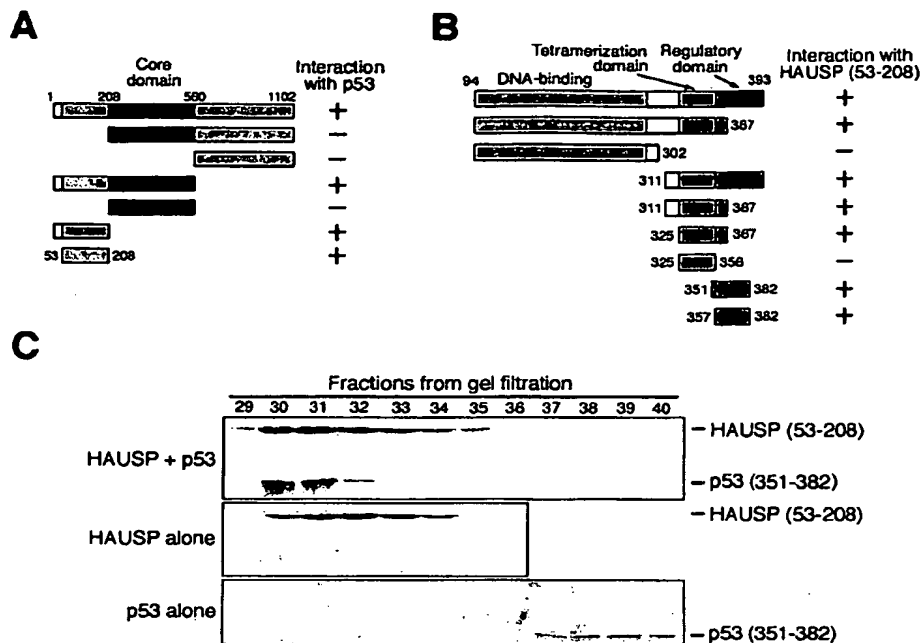


Figure 6. The N-Terminal Domain of HAUSP Recognizes the C-Terminal Peptide Sequences of p53

(A) Identification of the N-terminal domain of HAUSP as the primary p53 binding motif. Various purified recombinant HAUSP fragments were individually incubated with a large p53 fragment (residues 94–393) and then subjected to a gel filtration analysis. The results are summarized here.

(B) Mapping of a minimal p53 fragment that is both necessary and sufficient for the formation of a stable complex with HAUSP.

(C) A representative gel filtration run for the complex between HAUSP (53–208) and p53 (351–382). Aliquots of the fractions were visualized by Coomassie staining following SDS-PAGE. The p53 (351–382) peptide, which in isolation was eluted from gel filtration in fractions 38–39, was eluted in fraction 31–33 when complexed with HAUSP (53–208).

in the HAUSP catalytic core domain upon Ubal binding. In contrast to the Ubal binding-regulated enzyme activity for HAUSP, the  $\mu$  calpain protease is regulated by the local  $\text{Ca}^{2+}$  concentration and only acts upon surrounding proteins in response to elevated  $\text{Ca}^{2+}$ .

Polyubiquitin conjugated to target proteins can be assembled through different lysine residues of ubiquitin, producing different outcomes (Pickart, 2000). Polyubiquitin linked by Lys48 or, more rarely, Lys29 is thought to direct proteins to the 26S proteasome for degradation, whereas Lys63-linked polyubiquitin appears to signal other responses such as DNA repair and endocytosis. Both Lys29 and Lys63 are largely solvent accessible in the ubiquitin moiety of the HAUSP-Ubal complex, but the Lys48 side chain is involved in hydrogen bonds to residues in helix  $\alpha 5$  of HAUSP and is only partially exposed to solvent. Examination of the local structure indicates that the side chain of Lys48 can be linked to the C-terminal carboxylate of another ubiquitin without disruption of the HAUSP-Ubal interface. Nevertheless, Lys48 is not as freely available as the other ubiquitin lysine residues. Moreover, the HAUSP residues Asp305 and Glu308 that coordinate ubiquitin Lys48 are conserved as acidic amino acids in some other UBPs (e.g., scUBP15 and dFAF in Figure 2). A possible function of these residues is to promote binding of a ubiquitin moiety that contains a free Lys48 side chain. Thus, cleavage

by HAUSP might be biased toward the ubiquitin at the distal end of a K48-linked chain or toward conjugates with either monoubiquitin or non-K48-linked polyubiquitin.

At present, which ubiquitin lysine(s) is used in the MDM2-mediated polyubiquitination of p53 has not been established. It also is unclear whether the p53 conjugates acted on by HAUSP contain one or more polyubiquitin chains or instead are monoubiquitinated at multiple sites. Several lysines clustered near the C terminus function as ubiquitination sites on p53 (Nakamura et al., 2000; Rodriguez et al., 2000), and simultaneous mutation of these lysine residues (lysine 370, 372, 373, 381, 382, and 386) interfered with MDM2-dependent ubiquitination and degradation of p53. Surprisingly, ubiquitination of p53 by MDM2 in vitro was reported to yield conjugates that contained multiple mono-ubiquitin moieties (Lai et al., 2001). Establishing the specificity of HAUSP and whether it can remove entire polyubiquitin chains or only monoubiquitins from p53 should give considerable insight into the nature of ubiquitin-p53 conjugates in vivo. The presence of the p53 binding domain at the N-terminal portion of HAUSP argues strongly that it may prefer to cleave a conjugate at the proximal ubiquitin, i.e., between p53 and the attached mono- or polyubiquitin chain. In our in vitro assays, full-length HAUSP and the catalytic core domain were equally able to cleave

diubiquitin into mono-ubiquitin (Figure 4E and data not shown), and the full-length HAUSP removed all of the ubiquitins from the p53 conjugates (Li et al., 2002 and data not shown). However, this point, as well as an evaluation of HAUSP deubiquitination activity on mono-ubiquitin versus polyubiquitin p53 conjugates, awaits quantitative kinetic experiments with defined substrates.

In this study, we identified the C-terminal peptide sequences of p53 as necessary and sufficient for binding to the p53-recognition domain of HAUSP. The minimal peptide (residues 357–382) comprises only 26 amino acids (Figure 6B) but contains five of the six putative ubiquitination sites. The extensive overlap of these multiple ubiquitination sites with the HAUSP binding site imposes important structural constraints on HAUSP. First, the HAUSP p53-recognition domain must be in close proximity to the active site cleft of the catalytic domain. Second, to accommodate these multiple ubiquitin attachment sites, the connection between the HAUSP catalytic core domain and the p53 binding domain must be flexible in order to position each ubiquitin-p53 linkage within the active site for nucleophilic attack by Cys223. This organization is likely to be critical for the efficient deubiquitination of p53 *in vivo*.

#### Experimental Procedures

##### Protein Preparation

All constructs were generated using a standard PCR-based cloning strategy. The catalytic core domain of HAUSP (208–560) and all mutants were cloned into the vector pGEX-2T (Pharmacia). The full-length HAUSP protein was cloned into the BaculoGold virus (PharMingen) and expressed in insect cells. All HAUSP proteins were purified using glutathione sepharose 4B resin as described (Chai et al., 2001). The GST moiety was proteolytically removed by thrombin for all recombinant proteins in this study. Seleno-Met-substituted HAUSP (208–560) was generated as described (Qin et al., 1999).

##### Generation of Ubiquitin Aldehyde (Ubal) and a HAUSP-Ubal Complex

Ubal was prepared by carboxypeptidase Y-catalyzed exchange of 3-amino-1, 2-propanediol for ubiquitin Gly76 and the subsequent oxidation of the ubiquitin-diol product with NaIO<sub>4</sub> (Dunten and Cohen, 1989; Lam et al., 1997b). The Ubal thus obtained was incubated in 4-fold excess over HAUSP protein at pH 8 (25 mM Tris, 100 mM NaCl, and 5 mM DTT), and the HAUSP-Ubal complex was isolated by gel filtration (Superdex 200, 10 mM Tris, [pH 8.0], 100 mM NaCl, and 4 mM DTT).

##### Crystallization and Data Collection for HAUSP

Crystals were grown by the hanging-drop method by mixing the HAUSP protein (residues 208–560) (~15 mg/ml) with an equal volume of reservoir solution containing 100 mM Tris, [pH 7.0], and 20% PEG1000 (w/v). Small crystals appeared overnight and were used as seeds to generate larger crystals from Seleno-Met HAUSP protein. The crystals belong to the spacegroup P2<sub>1</sub>, with  $a = 75.75$  Å,  $b = 68.74$  Å,  $c = 76.34$  Å, and  $\beta = 95.4^\circ$ . There are two molecules per asymmetric unit. Crystals were equilibrated in a cryoprotectant buffer containing reservoir buffer plus 20% glycerol (v/v) and were flash frozen in a cold nitrogen stream at  $-170^\circ\text{C}$ . The native and MAD data set was collected at NSLS beamlines X-25 and X-12C, respectively, and processed using the software Denzo and Scalepack (Otwinowski and Minor, 1997).

##### Structure Determination

The structure was determined by multiple anomalous dispersion (Table 1). Data collected at three wavelengths were treated with

SOLVE (Terwilliger and Berendzen, 1996) and 18 selenium atoms per asymmetric unit were located. These selenium positions were further refined using MLPHARE (CCP4, 1994). Initial MAD phases, with a mean figure of merit of 0.63 at 2.6 Å resolution, were extended to 2.3 Å and improved with solvent flattening and histogram matching using DM (CCP4, 1994). A model was built using O (Jones et al., 1991) and refined at 2.3 Å resolution using CNS (Brunger et al., 1998). The final refined model contains two molecules. One molecule contains residues 208–410, 418–501, and 509–554. The other molecule contains residues 208–410, 418–501, and 509–555.

##### Crystallization and Structure Determination of the HAUSP-Ubal Complex

Crystals were grown by the hanging-drop method by mixing the complex (~10 mg/ml) with an equal volume of reservoir solution containing 100 mM citrate, [pH 5.5], and 20% PEG3000. The crystals belong to the spacegroup P2<sub>1</sub>2<sub>1</sub>2<sub>1</sub>, with  $a = 99.5$  Å,  $b = 101.2$  Å, and  $c = 141.3$  Å. The native data set was collected at NSLS X-25 and processed using the software Denzo and Scalepack (Otwinowski and Minor, 1997).

The structure was determined by molecular replacement, using the software AMoRe (Navaza, 1994). The coordinates of HAUSP were used for rotational research against all reflections between 15 and 3.0 Å in the native data set. The top 50 solutions from the rotational search were individually used for a subsequent translational search, which yielded two excellent solutions. This model was examined with the program O (Jones et al., 1991). Refinement by the program CNS (Brunger et al., 1998), against the 2.3 Å native data set, allowed visualization of an additional HAUSP fragment. A model was built with the program O and refined further by simulated annealing. The final atomic model contains two copies of the HAUSP-Ubal complex and an isolated HAUSP protein. Both complexes contain HAUSP residues 208–554 and Ubal residues 1–76. The isolated HAUSP protein contains residues 208–410 and 418–554.

##### In Vitro Deubiquitination Assays

The following ubiquitin conjugates were used as substrates: ubiquitin-7-amido-4-methylcoumarin (ubiquitin-AMC; from BostonBiochem); ubiquitin fused to the N terminus of unfolded and lucifer yellow (LY)-labeled chicken ovomucoid domain I (ubiquitin-OM; Yao and Cohen, 2002); K48-linked diubiquitin (Chen and Pickart, 1990), and ubiquitinated p53 (Li et al., 2002). Quantitative assays were done by incubation of substrate with HAUSP at 37°C in 50 mM [pH 8.0] HEPES, 50 mM NaCl, 5 mM DTT, 1 mM EDTA, and 0.1 mg/ml ovalbumin. Ubiquitin-AMC cleavage was monitored fluorometrically (Dang et al., 1998), and diubiquitin cleavage was determined by cation-exchange HPLC (Lam et al., 1997a). The disassembly of fluorescent LY-labeled ubiquitin-OM was visualized and quantified with a cooled CCD camera system (BioChem System, UVP Bioluminescence) after separation by SDS-PAGE. To survey the effects of various mutations in HAUSP, an equal amount (0.5 µg) of the wild-type or mutant HAUSP proteins (208–560) was incubated with 2 µg of the diubiquitin substrate at 37°C in 30 µl reaction buffer containing 25 mM Tris, [pH 8.0], 100 mM NaCl, 100 µg/ml BSA, and 2 mM DTT. The reaction was stopped by the addition of 25 µl 2× SDS sample buffer and analyzed by SDS PAGE.

##### An Interaction Assay by Gel Filtration

Gel filtration was employed to examine the interaction between p53 and HAUSP. The details were as described (Wu et al., 2001).

##### Acknowledgments

We thank A. Saxena at NSLS-X12C and M. Becker at NSLS-X25 for help, and N. Hunt for administrative assistance.

Received: October 18, 2002

Revised: November 21, 2002

## References

- Baek, K.-H., Mondoux, M.A., Jaster, R., Fire-Levin, E., and D'Andrea, A.D. (2001). DUB-2A, a new member of the DUB subfamily of hematopoietic deubiquitinating enzymes. *Blood* 98, 636-642.
- Brunger, A.T., Adams, P.D., Clore, G.M., Delano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., et al. (1998). Crystallography and NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr. D* 54, 905-921.
- Chai, J., Shiozaki, E., Srinivasula, S.M., Wu, Q., Datta, P., Alnemri, E.S., and Shi, Y. (2001). Structural basis of caspase-7 inhibition by XIAP. *Cell* 104, 769-780.
- Chen, Z., and Pickart, C.M. (1990). A 25-kilodalton ubiquitin carrier protein (E2) catalyzes multi-ubiquitin chain synthesis via lysine 48 of ubiquitin. *J. Biol. Chem.* 265, 21835-21842.
- Chung, C.H., and Baek, S.H. (1999). Deubiquitinating enzymes: their diversity and emerging roles. *Biochem. Biophys. Res. Commun.* 266, 633-640.
- Chung, J.Y., Park, Y.C., Ye, H., and Wu, H. (2002). All TRAFs are not created equal: common and distinct molecular mechanisms of TRAF-mediated signal transduction. *J. Cell Sci.* 115, 679-688.
- Chung, K.K., Dawson, V.L., and Dawson, T.M. (2001). The role of the ubiquitin-proteasomal pathway in Parkinson's disease and other neurodegenerative disorders. *Trends Neurosci.* 24, S7-S14.
- CCP4 (Collaborative Computational Project 4) (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* 50, 760-763.
- Conaway, R.C., Brower, C.S., and Conaway, J.W. (2002). Emerging roles of ubiquitin in transcription regulation. *Science* 296, 1254-1258.
- D'Andrea, A., and Pellman, D. (1998). Deubiquitinating enzymes: a new class of biological regulators. *Crit. Rev. Biochem. Mol. Biol.* 33, 337-352.
- Dang, L.C., Melandri, F.D., and Stein, R.L. (1998). Kinetic and mechanistic studies on the hydrolysis of ubiquitin C-terminal 7-amino-4-methylcoumarin by deubiquitinating enzymes. *Biochemistry* 37, 1868-1879.
- Dunten, R.L., and Cohen, R.E. (1989). Recognition of modified forms of ribonuclease A by the ubiquitin system. *J. Biol. Chem.* 264, 16739-16747.
- Everett, R.D., Meredith, M., Orr, A., Cross, A., Kathoria, M., and Parkinson, J. (1997). A novel ubiquitin-specific protease is dynamically associated with the PML nuclear domain and binds to a herpesvirus regulatory protein. *EMBO J.* 16, 566-577.
- Gilchrist, C.A., and Baker, R.T. (2000). Characterization of the ubiquitin-specific protease activity of the mouse/human Unp/Unph oncoprotein. *Biochim. Biophys. Acta* 1481, 297-309.
- Glickman, M.H., and Ciechanover, A. (2002). The ubiquitin-proteasome proteolytic pathway: destruction for the sake of construction. *Physiol. Rev.* 82, 373-428.
- Henikoff, S., Henikoff, J.G., Alford, W.J., and Pietrzkowski, S. (1995). Automated construction and graphical presentation of protein blocks from unaligned sequences. *Gene* 163, GC17-26.
- Hershko, A., Ciechanover, A., and Varshavsky, A. (2000). The ubiquitin system. *Nat. Med.* 6, 1073-1081.
- Hershko, A., and Rose, I.A. (1987). Ubiquitin-aldehyde: a general inhibitor of ubiquitin-recycling processes. *Proc. Natl. Acad. Sci. USA* 84, 1829-1833.
- Hochstrasser, M. (1996). Ubiquitin-dependent protein degradation. *Annu. Rev. Genet.* 30, 405-439.
- Hochstrasser, M. (1998). There's the rub: a novel ubiquitin-like modification linked to cell cycle regulation. *Genes Dev.* 12, 901-907.
- Holm, L., and Sander, C. (1993). Protein structure comparison by alignment of distance matrices. *J. Mol. Biol.* 233, 123-138.
- Huang, Y., Baker, R.T., and Fischer-Vize, J.A. (1995). Control of cell fate by a deubiquitinating enzyme encoded by the *fat facets* gene. *Science* 270, 1828-1831.
- Johnston, S.C., Larsen, C.N., Cook, W.J., Wilkinson, K.D., and Hill, C.P. (1997). Crystal structure of a deubiquitinating enzyme (human UCH-L3) at 1.8 Å resolution. *EMBO J.* 16, 3787-3796.
- Johnston, S.C., Riddle, S.M., Cohen, R.E., and Hill, C.P. (1999). Structural basis for the specificity of ubiquitin C-terminal hydrolysis. *EMBO J.* 18, 3877-3887.
- Jones, T.A., Zou, J.-Y., Cowan, S.W., and Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A* 47, 110-119.
- Kraulis, P.J. (1991). Molscript: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* 24, 946-950.
- Lai, Z., Ferry, K.V., Diamond, M.A., Wee, K., Kim, Y., Ma, J., Yang, T., Benfield, P.A., Copeland, R.A., and Auger, K.R. (2001). Human mdm2 mediates multiple mono-ubiquitination of p53 by a mechanism requiring enzyme isomerization. *J. Biol. Chem.* 276, 31357-31367.
- Lam, Y.A., DeMartino, G.N., Pickart, C.M., and Cohen, R.E. (1997a). Specificity of the ubiquitin isopeptidase in the PA700 regulatory complex of 26 S proteasomes. *J. Biol. Chem.* 272, 28438-28446.
- Lam, Y.A., Xu, W., DeMartino, G.N., and Cohen, R.E. (1997b). Editing of ubiquitin conjugates by an isopeptidase in the 26S proteasome. *Nature* 385, 737-740.
- Laney, J.D., and Hochstrasser, M. (1999). Substrate targeting in the ubiquitin system. *Cell* 97, 427-430.
- Larsen, C.N., Price, J.S., and Wilkinson, K.D. (1996). Substrate binding and catalysis by ubiquitin C-terminal hydrolases: identification of two active site residues. *Biochemistry* 35, 6735-6744.
- Levine, A.J. (1997). p53, the cellular gatekeeper for growth and division. *Cell* 88, 323-331.
- Li, M., Chen, D., Shiloh, A., Luo, J., Nikolaev, A.Y., and Gu, W. (2002). Deubiquitination of p53 by HAUSP is an important pathway for p53 stabilization. *Nature* 416, 638-653.
- Moldoveanu, T., Hosfield, C.M., Lim, D., Elice, J.S., Jia, Z., and Davies, P.L. (2002). A Ca<sup>2+</sup> switch aligns the active site of calpain. *Cell* 108, 649-660.
- Mossessova, E., and Lima, C.D. (2000). Ulp1-SUMO crystal structure and genetic analysis reveal conserved interactions and a regulatory element essential for cell growth in yeast. *Mol. Cell* 5, 865-876.
- Nakamura, S., Roth, J.A., and Mukhopadhyay, T. (2000). Multiple lysine mutations in the C-terminal domain of p53 interfere with MDM2-dependent protein degradation and ubiquitination. *Mol. Cell. Biol.* 20, 9391-9398.
- Navaza, J. (1994). AMoRe and automated package for molecular replacement. *Acta Crystallogr. A* 50, 157-163.
- Nicholls, A., Sharp, K.A., and Honig, B. (1991). Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins* 11, 281-296.
- Otwinowski, Z., and Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* 276, 307-326.
- Pickart, C.M. (2000). Ubiquitin in chains. *Trends Biochem. Sci.* 25, 544-548.
- Pickart, C.M. (2001). Mechanisms underlying ubiquitination. *Annu. Rev. Biochem.* 70, 503-533.
- Qin, H., Srinivasula, S.M., Wu, G., Fernandes-Alnemri, T., Alnemri, E.S., and Shi, Y. (1999). Structural basis of procaspase-9 recruitment by the apoptotic protease-activating factor 1. *Nature* 399, 547-555.
- Rawlings, N.D., and Barrett, A.J. (1994). Families of cysteine peptidases. *Methods Enzymol.* 244, 461-488.
- Rodríguez, M.S., Desterro, J.M.P., Lain, S., Lane, D.P., and Hay, R. (2000). Multiple C-terminal lysine residues target p53 for ubiquitin-proteasome-mediated degradation. *Mol. Cell. Biol.* 20, 8458-8467.
- Schwartz, A.L., and Ciechanover, A. (1999). The ubiquitin-proteasome pathway and pathogenesis of human diseases. *Annu. Rev. Med.* 50, 57-74.
- Steitz, T.A. (1999). DNA polymerases: structural diversity and common mechanisms. *J. Biol. Chem.* 274, 17395-17398.

Terwilliger, T.C., and Berendzen, J. (1996). Correlated phasing of multiple isomorphous replacement data. *Acta Crystallogr. D* 52, 749–757.

Thompson, J.D., Higgins, D.G., and Gibson, T.J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22, 4673–4680.

Vogelstein, B., Lane, D., and Levine, A.J. (2000). Surfing the p53 network. *Nature* 408, 307–310.

Wilkinson, K.D. (1997). Regulation of ubiquitin-dependent processes by deubiquitinating enzymes. *FASEB J.* 11, 1245–1256.

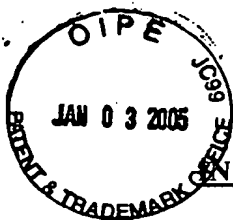
Wu, J.-W., Falman, R., Penry, J., and Shi, Y. (2001). Formation of a stable heterodimer between Smad2 and Smad4. *J. Biol. Chem.* 276, 20688–20694.

Yao, T., and Cohen, R.E. (2002). A cryptic protease couples deubiquitination and degradation by the proteasome. *Nature* 419, 403–407.

Zapata, J.M., Pawlowski, K., Haas, E., Ware, C.F., Godzik, A., and Reed, J.C. (2001). A diverse family of proteins containing tumor necrosis factor receptor-associated factor domains. *J. Biol. Chem.* 276, 24242–24252.

#### Accession Numbers

Atomic coordinates have been deposited with the Protein Data Bank (accession number 1NB8 for HAUSP and 1NBF for the HAUSP-Ubal complex).



patent TM\$

**THE UNITED STATES PATENT AND TRADEMARK OFFICE**

Commissioner for Patents  
P.O. Box 1450  
Alexandria, VA 22313-1450

Our file: 7616/99/129  
Applicant: Yigong Shi  
Serial No.:  
Filing Date:  
Title: Method of Screening Compounds that Antagonize HAUSP

Sir:  
Enclosed for filing in the United States Patent and Trademark Office is the following:

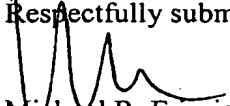
1. Provisional Patent Application (43 Pages)
2. Provisional Application for Patent Cover Sheet
3. Check No. 8178 for \$100.00
4. Transmittal Sheet
5. Postcard Receipt

**CONDITIONAL PETITION**


If any extension of time is required for the submission of the above-identified items, Applicant requests that this be considered a petition therefor. Please charge any additional charges or any other charges relating to this matter to deposit account of the writer, **Account No. 06-2143**. A duplicate copy of this letter is enclosed.

12/30/04  
Date

Respectfully submitted,

  
Michael R. Friscia  
Registration No. 33,884  
Wolff & Samson PC  
One Boland Drive  
West Orange, NJ 07052  
Tel: (973) 530-2028  
Fax: (973) 530-2288

I hereby certify that this correspondence is being deposited with the United States Postal Service, postage prepaid, as "Express Mail Post Office to Addressee," Mailing Label No. **EV335727905US**, to the Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450 on 12/30/04.

By   
Janelle Fava



01-03-2005

U.S. Patent & TMO/TM Mail Rpt DL #39

EV335727905W

Please type a plus sign (+) inside this box 

PTO/SB/16 (12-04)

Approved for use through 07/31/2008. OMB 0651-0032  
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

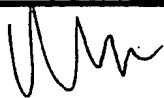
**PROVISIONAL APPLICATION FOR PATENT COVER SHEET**

This is a request for filing a PROVISIONAL APPLICATION FOR PATENT under 37 CFR 1.53(c).

INVENTOR(S)					
Given Name (first and middle (if any))		Family Name or Surname		Residence (City and either State or Foreign Country)	
Yigong		Shi		Plainsboro, NJ	
<input type="checkbox"/> Additional inventors are being named on the _____ separately numbered sheets attached hereto					
TITLE OF THE INVENTION (280 characters max)					
Method of Screening Compounds that Antagonize HAUSP					
Direct all correspondence to: CORRESPONDENCE ADDRESS					
<input checked="" type="checkbox"/> Customer Number		20694		<div>Place Customer Number Bar Code Label here</div>	
OR Type Customer Number here					
<input type="checkbox"/> Firm or Individual Name					
Address					
Address					
City		State		ZIP	
Country		Telephone		Fax	
ENCLOSED APPLICATION PARTS (check all that apply)					
<input checked="" type="checkbox"/>	Specification	Number of Pages	43	<input type="checkbox"/>	CD(s), Number
<input type="checkbox"/>	Drawing(s)	Number of Sheets		<input type="checkbox"/>	Other (specify)
<input type="checkbox"/>	Application Data Sheet. See 37 CFR 1.76				
Total # of sheets		43		Application Size Fee = \$0.00	
METHOD OF PAYMENT OF FILING FEES FOR THIS PROVISIONAL APPLICATION FOR PATENT (check one)					
<input type="checkbox"/>	Applicant claims small entity status. See 37 CFR 1.27.				FILING FEE AMOUNT (\$)
<input checked="" type="checkbox"/>	A check or money order is enclosed to cover the filing fees				\$100.00
<input type="checkbox"/>	The Director is hereby authorized to charge filing fees or credit any overpayment to Deposit Account Number				
<input type="checkbox"/>	Payment by credit card. Form PTO-2038 is attached.				
The invention was made by an agency of the United States Government or under a contract with an agency of the United States Government.					
<input type="checkbox"/>	No.				
<input checked="" type="checkbox"/>	Yes, the name of the U.S. Government agency and the Government contract number are: NIH 126-6087				

Respectfully submitted,

SIGNATURE



Date

12/30/04

TYPED or PRINTED NAME

Michael R. Friscia

REGISTRATION NO.

33,884

(if appropriate)

Docket Number:

7616/99/129

TELEPHONE

(973) 530-2024

**USE ONLY FOR FILING A PROVISIONAL APPLICATION FOR PATENT**

This collection of information is required by 37 CFR 1.51. The information is used by the public to file (and by the PTO to process) a provisional application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 8 hours to complete, including gathering, preparing, and submitting the complete provisional application to the PTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

P19SMALL/REV07